

Finite-Length Scaling of Polar Codes

S. Hamed Hassani, Kasra Alishahi, and Rüdiger Urbanke

Abstract—We consider the tradeoff between the rate and the block-length for a fixed error probability when we use polar codes and the successive cancellation decoder. The “scaling” between these two parameters gives interesting engineering insights, and in particular tells us how fast one can approach capacity if our goal is to achieve a fixed block-error probability.

Noticing the fact that for polar codes the exact scaling behavior is greatly dependent on the choice of the channel, our objective is to provide scaling laws that hold universally for all the BMS channels. Our approach is based on analyzing the dynamics of the un-polarized channels. More precisely, we provide bounds on (the exponent of) the number of sub-channels whose Bhattacharyya constant falls in a fixed interval $[a, b]$. Mathematically, this can be stated as bounding the sequence $\{\frac{1}{n} \log \Pr(Z_n \in [a, b])\}_{n \in \mathbb{N}}$, where Z_n is the Bhattacharyya process. We then use these bounds to derive trade-offs between the rate and the block-length.

The main results of this paper can be summarized as follows. Let W be a BMS channel with capacity $I(W)$. Consider the sum of Bhattacharyya parameters of sub-channels chosen (by the polar coding scheme) to transmit information. If we require this sum to be smaller than a given value $P_e > 0$, then the required block-length N scales in terms of the rate $R < I(W)$ as $N \geq \frac{\alpha}{(I(W)-R)^\mu}$, where α is a positive constant that depends on P_e and $I(W)$. We show that $\mu = 3.55$ is a valid choice, and we conjecture that indeed the value of μ can be improved to $\underline{\mu} = 3.627$, the parameter for the binary erasure channel. Also, we show that with the same requirement on the sum of Bhattacharyya parameters, the block-length scales in terms of the rate like $N \leq \frac{\beta}{(I(W)-R)^\mu}$, where β is a constant that depends on P_e and $I(W)$, and $\bar{\mu} = 7$.

I. INTRODUCTION

Polar coding schemes [1] provably achieve the capacity of a wide array of channels including binary memoryless symmetric (BMS) channels.

In coding, the three most important parameters are: rate (R), block-length (N), and block error probability (P_e). Ideally, given a family of codes such as the family of polar codes, one would like to be able to describe the exact relationship between these three parameters. This however is a formidable task. Slightly easier is to fix one of the parameters and then to describe the relationship (scaling) of the remaining two.

E.g., assume that we fix the rate and consider the relationship between the error probability and the block-length. This is the study of the classical error exponent. For instance, for random codes a closer look shows that $P_e = e^{-NE(R,W)+o(N)}$, where $E(R, W)$ is the so-called *random error exponent* [2] of the channel W . For polar codes, Arikan and Telatar [3] showed that when W is a BMS channel, for any rate $R < I(W)$ the block error probability is upper bounded by 2^{-N^β} for any

$\beta < \frac{1}{2}$ and N large enough. This result was refined later in [4] to be dependent on R , i.e. for polar codes with the successive cancellation (SC) decoder

$$P_e = 2^{-2^{\frac{n}{2} + \sqrt{n}Q^{-1}(\frac{R}{I(W)}) + o(n)}},$$

where¹ $n = \log N$ and $Q(t) \triangleq \int_t^\infty e^{-z^2/2} dz / \sqrt{2\pi}$.

Another option is to fix the error probability and to consider the relationship between the block-length and the rate. In other words, given a code and a desired (and fixed) error probability P_e , what is the block-length N required, in terms of the rate R , so that the code has error probability less than P_e ? This scaling is arguably more relevant (than the error exponent) from a practical point of view since we typically have a certain requirement on the error probability and then are interested in using the shortest code possible to transmit at a certain rate.

As a benchmark, let us mention what is the shortest block-length that we can hope for. Some thought clarifies that the random variations of the channel itself require $R \leq I(W) - \Theta(\frac{1}{\sqrt{N}})$ or equivalently $N \geq \Theta(\frac{1}{(I(W)-R)^2})$. Indeed, a sequence of works starting from [5], then [6], and finally [7] showed that the minimum possible block-length N required to achieve a rate R with a fixed error probability P_e is roughly equal to

$$N \approx \frac{V(Q^{-1}(P_e))^2}{(I(W) - R)^2}, \quad (1)$$

where V is a characteristic of the channel referred to as channel dispersion. In other words, the best codes require a block-length of order $\Theta(\frac{1}{(I(W)-R)^2})$.

The main objective of this paper is to characterize similar type of relations for polar codes with the SC decoder. We argue in this paper that this problem is fundamentally related to the dynamics of channel polarization and specially the speed of which the polarization is taking place. To state things in a more convenient language, let us start with some preliminary definitions and settings regarding polarization and polar codes.

A. Preliminaries

Let $W : \mathcal{X} \rightarrow \mathcal{Y}$ be a BMS channel, with input alphabet $\mathcal{X} = \{0, 1\}$, output alphabet \mathcal{Y} , and the transition probabilities $\{W(y|x) : x \in \mathcal{X}, y \in \mathcal{Y}\}$. We consider the following three parameters for the channel W ,

$$H(W) = \sum_{y \in \mathcal{Y}} W(y|1) \log \frac{W(y|1) + W(y|0)}{W(y|1)}, \quad (2)$$

$$Z(W) = \sum_{y \in \mathcal{Y}} \sqrt{W(y|0)W(y|1)}, \quad (3)$$

$$E(W) = \frac{1}{2} \sum_{y \in \mathcal{Y}} W(y|1) e^{-\frac{1}{2}(\ln \frac{W(y|1)}{W(y|0)} + |\ln \frac{W(y|1)}{W(y|0)}|)}. \quad (4)$$

¹In this paper all the logarithms are in base 2.

S. H. Hassani and R. Urbanke are with the School of Computer and Communication Science, EPFL, CH-1015 Lausanne, Switzerland (e-mail: {seyehamed.hassani, rudiger.urbanke}@epfl.ch).

K. Alishahi is with the department of Mathematical Sciences, Sharif University of Technology, Tehran, Iran (email: alishahi@sharif.edu).

The parameter $H(W)$ is equal to the entropy of the output of W given its input when we assume uniform distribution on the inputs, i.e., $H(W) = H(X|Y)$. Hence, we call the parameter $H(W)$ the entropy of the channel W . Also note that the capacity of W , which we denote by $I(W)$, is given by $I(W) = 1 - H(W)$. The parameter $Z(W)$ is called the Bhattacharyya parameter of W and $E(W)$ is called the error probability of W . It can be shown that $E(W)$ is equal to the error probability in estimating the channel input x on the basis of the channel output y via the maximum-likelihood decoding of $W(y|x)$ (with the further assumption that the input has uniform distribution). The following relations hold between these parameters (see for e.g., [1] and [13, Chapter 4]):

$$0 \leq 2E(W) \leq H(W) \leq Z(W) \leq 1, \quad (5)$$

$$H(W) \leq h_2(E(W)), \quad (6)$$

$$Z(W) \leq \sqrt{1 - (1 - H(W))^2}, \quad (7)$$

where $h_2(\cdot)$ denotes the binary entropy function, i.e.,

$$h_2(x) = -x \log_2(x) - (1-x) \log_2(1-x). \quad (8)$$

B. Channel transform

Let \mathcal{W} denote the set of all the BMS channels and consider a transform $W \rightarrow (W^0, W^1)$ that maps \mathcal{W} to \mathcal{W}^2 in the following manner. Having the channel $W : \{0, 1\} \rightarrow \mathcal{Y}$, the channels $W^0 : \{0, 1\} \rightarrow \mathcal{Y}^2$ and $W^1 : \{0, 1\} \rightarrow \{0, 1\} \times \mathcal{Y}^2$ are defined as

$$W^0(y_1, y_2|x_1) = \sum_{x_2 \in \{0, 1\}} \frac{1}{2} W(y_1|x_1 \oplus x_2) W(y_2|x_2) \quad (9)$$

$$W^1(y_1, y_2, x_1|x_2) = \frac{1}{2} W(y_1|x_1 \oplus x_2) W(y_2|x_2), \quad (10)$$

A direct consequence of the chain rule of entropy yields

$$\frac{H(W^0) + H(W^1)}{2} = H(W). \quad (11)$$

Regarding the other parameters, we have

$$Z(W) \leq Z(W^0) \leq 1 - (1 - Z(W))^2, \quad (12)$$

$$Z(W^1) = Z(W)^2, \quad (13)$$

and

$$E(W^0) = 1 - (1 - E(W))^2, \quad (14)$$

$$E(W)^2 \leq E(W^1) \leq E(W). \quad (15)$$

C. Channel Polarization

Consider an infinite binary tree with the root node placed at the top. In this tree each vertex has 2 children and there are 2^n vertices at level n . Assume that we label these vertices from left to right from 0 to $2^n - 1$. Here, we intend to assign to each vertex of the tree a BMS channel. We do this by a recursive procedure. Assign to the root node the channel W itself. Now consider the channel splitting transform $W \rightarrow (W^0, W^1)$ and from left to right, assign W^0 and W^1 to the children of the root node. In general, if Q is the channel that is assigned to vertex v , we assign Q^0 and Q^1 , from left to right respectively, to the children of the node v . In this way, we recursively assign a

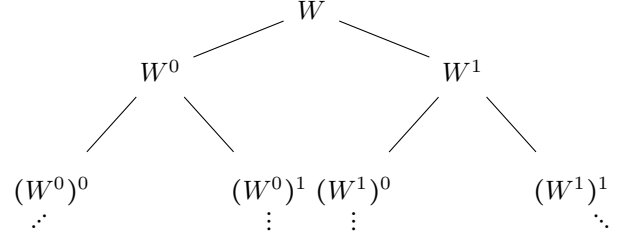


Fig. 1. The infinite ℓ -ary tree and the channels assigned to it for $\ell = 2$.

channel to all the vertices of the tree. Figure 1 shows the first 2 levels of the binary tree. Assuming $N = 2^n$, we let $W_N^{(i)}$ denote the channel that is assigned to a vertex with label i at level n of the tree, $0 \leq i \leq N - 1$. As a result, one can equivalently relate the channel $W_N^{(i)}$ to W via the following procedure: let the binary representation of i be $b_1 b_2 \dots b_n$, where b_1 is the most significant digit. Then we have

$$W_N^{(i)} = (((W^{b_1})^{b_2}) \dots)^{b_n}.$$

As an example, assuming $i = 6$, $n = 3$ we have $W_8^{(7)} = ((W^1)^1)^0$. We now proceed with defining a stochastic process called the polarization process. This process can be considered as a stochastic representation of the channels associated to different levels of the infinite binary tree.

D. Polarization Process

Let $\{B_n, n \geq 1\}$ be a sequence of iid Bernoulli($\frac{1}{2}$) random variables. Denote by $(\mathcal{F}, \Omega, \Pr)$ the probability space generated by this sequence and let $(\mathcal{F}_n, \Omega_n, \Pr_n)$ be the probability space generated by (B_1, \dots, B_n) . For a BMS channel W , define a random sequence of channels W_n , $n \in \mathbb{N} \triangleq \{0, 1, 2, \dots\}$, as $W_0 = W$ and

$$W_n = \begin{cases} W_{n-1}^0 & \text{if } B_n = 0, \\ W_{n-1}^1 & \text{if } B_n = 1, \end{cases} \quad (16)$$

where the channels on the right side are given by the transform $W_{n-1} \rightarrow (W_{n-1}^0, W_{n-1}^1)$. Let us also define the random processes $\{H_n\}_{n \in \mathbb{N}}$, $\{I_n\}_{n \in \mathbb{N}}$, $\{Z_n\}_{n \in \mathbb{N}}$ and $\{E_n\}_{n \in \mathbb{N}}$ as $H_n = H(W_n)$, $I_n = I(W_n) = 1 - H(W_n)$, $Z_n = Z(W_n)$ and $E_n = E(W_n)$.

Example 1: By a straightforward calculation one can show that for $W = \text{BEC}(z)$ we have

$$W^0 = \text{BEC}(1 - (1 - z)^2) \quad (17)$$

$$W^1 = \text{BEC}(z^2). \quad (18)$$

Hence, when $W = \text{BEC}(z)$, the channel W_n is always a BEC. Furthermore, the processes H_n, I_n, Z_n and E_n admit simple closed form recursions as follows. We have $H_0 = z$ and for $n \geq 1$

$$H_n = \begin{cases} 1 - (1 - H_{n-1})^2, & \text{w.p. } \frac{1}{2} \\ H_{n-1}^2, & \text{w.p. } \frac{1}{2}. \end{cases} \quad (19)$$

Also, we have² $2E_n = H_n = 1 - I_n = Z_n$.

²For the channel $W = \text{BEC}(z)$, it is easy to show that $2E(W) = H(W) = Z(W) = z$.

For channels other than the BEC, the channel W_n gets quite complicated in the sense that the cardinality of the output alphabet of the channel W_n is doubly exponential in n (or exponential in N). Thus, tracking the exact outcome of W_n seems to be a difficult task (for more detail see [15], [16]). Instead, as we will see in the sequel, one can prove many interesting properties regarding the processes H_n, Z_n and E_n .

Let us quickly review the limiting properties of the above mentioned processes [1], [3]. From (11) and (16), one can write for $n \geq 1$

$$\mathbb{E}[H(W_n) | W_{n-1}] \stackrel{(16)}{=} \frac{H(W_{n-1}^0) + H(W_{n-1}^1)}{2} \stackrel{(11)}{=} H(W_{n-1}). \quad (20)$$

Hence, the process H_n is a martingale. Furthermore, since H_n is also bounded (5), by Doob's martingale convergence theorems, the process H_n converges in \mathcal{L}^1 (and almost surely) to a limit random variable H_∞ . As the convergence is in \mathcal{L}^1 , as $n \rightarrow \infty$ we have

$$\mathbb{E}[|H_n - H_{n-1}|] = \mathbb{E}[|H(W_n^0) - H(W_n)|] \rightarrow 0.$$

As a result, we must have that $H(W_n^0) - H(W_n)$ converges to 0 almost surely (a.s.). We now claim that for a channel P , in order to have $H(P^0) = H(P)$ we must have $H(P) = 0$ (i.e., P is the noiseless channel) or $H(P) = 1$ (i.e., P is the completely noisy channel). By this claim and the fact that H_n converges a.s. to H_∞ , we conclude that H_∞ take its values in the set $\{0, 1\}$. Also, as $\mathbb{E}[H_n] = \mathbb{E}[H_\infty] = H(W)$, we obtain

$$H_\infty = \begin{cases} 0 & \text{w.p. } 1 - H(W), \\ 1 & \text{w.p. } H(W). \end{cases} \quad (21)$$

It remains to prove the claim mentioned above. We use the so called *extremes of information combining* inequalities [13]. Let P be an arbitrary BMS channel. To simplify notation, let $h = H(P)$ and also let $\epsilon \in [0, \frac{1}{2}]$ be such that $h_2(\epsilon) = H(P)$. We have

$$h \leq \underbrace{H(\text{BSC}(\epsilon^0))}_{h_2(2\epsilon(1-\epsilon))} \leq H(P^0) \leq \underbrace{H(\text{BEC}(h^0))}_{1-(1-h)^2}, \quad (22)$$

$$\underbrace{H(\text{BEC}(h^1))}_{h^2} \leq H(P^1) \leq \underbrace{H(\text{BSC}(\epsilon^1))}_{2h-h_2(2\epsilon(1-\epsilon))} \leq h. \quad (23)$$

Now, to prove the claim, assume that P is such that $H(P^0) = H(P)$. Using (22) we obtain $H(\text{BSC}(h^0)) = H(P)$ or equivalently $h_2(2\epsilon(1-\epsilon)) = h_2(\epsilon)$. As a result, ϵ must be a solution of the equation $\epsilon = 2\epsilon(1-\epsilon)$ which yields $\epsilon = 0, \frac{1}{2}$. Also, as $H(P) = h_2(\epsilon)$, then $H(P)$ can either be 0 or 1 and hence the claim is justified. Using the bounds (5)-(7) it is clear that the processes Z_n and E_n converge a.s. to H_∞ and $\frac{1}{2}H_\infty$, respectively.

E. Polar Codes

Given the rate $R < I(W)$, polar coding is based on choosing a set of $2^n R$ rows of the matrix $G_n = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}^{\otimes n}$ to form a $2^n R \times 2^n$ matrix which is used as the generator matrix in the encoding procedure. The way this set is chosen is dependent on the channel W and is briefly explained as follows: Choose a subset of size NR from the set of channels $\{W_N^{(i)}\}_{0 \leq i \leq N-1}$

that have the least possible error probability (given in (4)) and choose the rows G_n with the same indices as these channels. E.g., if the channel $W_N^{(i)}$ is chosen, then the i -th row of G_n is selected. In the following, given N , we call the set of indices of NR channels with the least error probability, the set of good indices and denote it by $\mathcal{I}_{N,R}$. In the sequel, we will frequently use the term “the set of good indices” and $\mathcal{I}_{N,R}$ interchangeably.

It is proved in [1] that the block error probability of such polar coding scheme under SC decoding, denoted by P_e , is bounded from both sides by³

$$\max_{i \in \mathcal{I}_{N,R}} E(W_N^{(i)}) \leq P_e \leq \sum_{i \in \mathcal{I}_{N,R}} E(W_N^{(i)}). \quad (24)$$

We now briefly explain why such a code construction is reliable for any rate $R < I(W)$, provided that the block-length is large enough. Recall from Section I-D that the process $E_n = E(W_n)$ converges a.s. to a r.v. E_∞ such that $\Pr(E_\infty = 0) = 1 - H(W) = I(W)$. Hence, it is clear from the definition of the set good indices, $\mathcal{I}_{N,R}$, that the left side of (24) decays to 0 as n grows large. However, the story is not over yet as this is only a lower bound on P_e . Nonetheless, one can also show that the right side of (24) decays to 0. This was initially shown in [1] and later in [3] the authors showed that all of the three terms in (24) behave like $2^{-2^{\frac{n}{2} + o(n)}}$.

II. PROBLEM FORMULATION

As we have seen in the previous section, the processes H_n and Z_n polarize in the sense that they converges a.s. to a $\{0, 1\}$ valued r.v. H_∞ and Z_∞ . Here, we investigate the dynamics of polarization. We start by noting that at each time n there still exists a (small and in n vanishing) probability that the process Z_n (or H_n) takes a value far away from the endpoints of the unit interval (i.e., 0 and 1). Our primary objective is to study these small probabilities. More concretely, let $0 < a < b < 1$ be constants and consider the quantity $\Pr(Z_n \in [a, b])$. This quantity represents the fraction of sub-channels that are still un-polarized at time n . An important question is how fast the quantity $\Pr(Z_n \in [a, b])$ decays to zero. This question is intimately related to measuring the limiting properties of the sequence $\{\frac{1}{n} \log \Pr(Z_n \in [a, b])\}_{n \in \mathbb{N}}$.

Example 2: Assume $W = \text{BEC}(z)$. In this case the process Z_n has a simple closed form recursion as $Z_0 = z$ and

$$Z_{n+1} = \begin{cases} Z_n^2, & \text{w.p. } \frac{1}{2}, \\ 1 - (1 - Z_n)^2, & \text{w.p. } \frac{1}{2}. \end{cases} \quad (25)$$

Hence, it is straightforward to compute the value $\Pr(Z_n \in [a, b])$ numerically. Let $a = 1 - b = 0.1$. Figure 2 shows the value $\frac{1}{n} \log(\Pr(Z_n \in [a, b]))$ in terms of n for $z = 0.5, 0.6, 0.7$. This figure suggests that the sequence $\{\frac{1}{n} \log \Pr(Z_n \in [a, b])\}$ converges to a limiting value that is somewhere between -0.27 and -0.28 . Note that for different values of z , the limiting values are very close to each other. \diamond

For other BMS channels, the process Z_n does not have a simple closed form recursion as for the BEC, and hence we

³Note here that by (4) the error probability of a BMS channel is less than its Bhattacharyya value. Hence, the right side of (24) is a better upper bound for the block error probability than the sum of Bhattacharyya values.

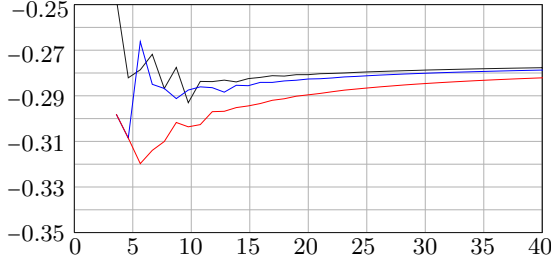


Fig. 2. The value of $\frac{1}{n} \log(\Pr(Z_n \in [a, b]))$ versus n for $a = 1 - b = 0.1$ when W is a BEC with erasure probability $z = 0.5$ (top curve), $z = 0.6$ (middle curve) and $z = 0.7$ (bottom curve).

need to use approximation methods (for more details see [15], [16]). Using such methods, we have plotted in Figure 3 the value of $\Pr(Z_n \in [a, b])$ ($a = 1 - b = 0.1$) for the channel families BSC(ϵ), and BAWGNC(σ) with different parameter values.

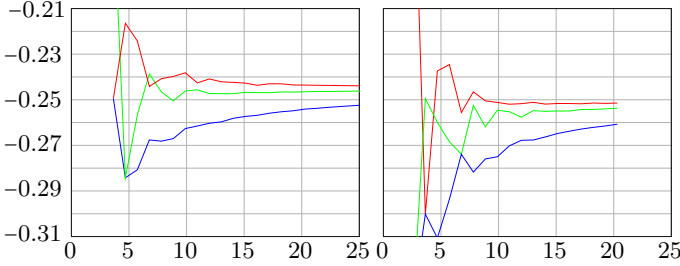


Fig. 3. *Left figure:* The value of $\frac{1}{n} \log(\Pr(Z_n \in [a, b]))$ versus n for $a = 1 - b = 0.1$ and W being a BSC with cross-over probability $\epsilon = 0.11, 0.146, 0.189$. These BSC channels have capacity 0.5, 0.4 and 0.3, respectively. *Right figure:* the value of $\frac{1}{n} \log(\Pr(Z_n \in [a, b]))$ versus n for $a = 1 - b = 0.1$ and W is a BAWGN with noise variance $\sigma = 0.978$ (top curve), $\sigma = 1.149$ (middle curve), and $\sigma = 1.386$ (bottom curve). These BAWGN channels have capacities 0.5, 0.4 and 0.3, respectively.

The above numerical evidence suggests that the quantity $\Pr(Z_n \in [a, b])$ decays to zero exponentially fast in n . Further, we observe that the limiting value of this sequence is dependent on the starting channel W (e.g., from the figures it is clear that the channels BEC, BSC and BAWGN have different limiting values). Let us now be concrete and rephrase the above speculations as follows.

Question 1: Does the quantity $\Pr(Z_n \in [a, b])$ decay exponentially in n ? If yes, what is the limiting value of $\frac{1}{n} \log \Pr(Z_n \in [a, b])$ and how is this limit related to the starting channel W and the choice of a and b ?

From Figures 2 and 3, we observe that the value of $\frac{1}{n} \log \Pr(Z_n \in [a, b])$ is the least when W is a BEC and this suggests that the channel BEC polarizes faster than the other BMS channels. This is intuitively justified as follows: Fix a value $z \in (0, 1)$ and assume that W is a BMS channel with Bhattacharyya parameter $Z(W) = z$. Now, consider the values $Z(W^0)$ and $Z(W^1)$. Using relations (12) and (13), it is clear that the values $Z(W^0)$ and $Z(W^1)$ are closest to the end points of the unit interval if W is a BEC. In other words, at the channel splitting transform, the channel BEC(z) polarizes faster than the other BMS channels.

Question 2: For which set of channels does the quantity

$\Pr(Z_n \in [a, b])$ decay the fastest or the slowest?

Let us now be more ambitious and aim for our ultimate goal.

Question 3: Can we characterize the exact behavior of $\Pr(Z_n \in [a, b])$ as a function of n, a, b and W ?

Finally, we ask how the answers to the above questions will guide us through the understanding of the finite-length scaling behavior of polar codes. An immediate relation stems from the fact that the quantity $\Pr(Z_n \in [a, b])$ indicates the portion of the sub-channels that have not polarized at time n . In particular, all the channels in this set have a large Bhattacharyya value and hence cannot be included in the set of good indices. Therefore, the maximum reliable rate that we can achieve is restricted by the portion of this yet un-polarized channels. Consequently, the answers to the above questions will be crucial in finding answers to the following question.

Question 4: Fix the channel W and a target block error probability P_e . To have a polar code with error probability less than P_e , how does the required block-length N scale with the rate R ?

Finding a suitable answer to the above questions is an easier task when the channel W is a BEC. This is due to the simple closed form expression of the process Z_n given in (25). In the next section (Section III), we provide heuristic methods that lead to suitable numerical answers to Questions 1 and 3 for the BEC. As we will see in the next section, such heuristic derivations are in excellent compliance with numerical experiments. Using such derivations, we also give an answer to Question 4 for the BEC. The heuristic results of Section III provide us then with a concrete path to analytically tackle the above questions. In Section IV we provide analytical answers to Questions 1-4 for the BEC as well as other BMS channels. Proving the full picture of Section III is beyond what we achieve in Section IV, nevertheless, we provide close and useful bounds.

III. HEURISTIC DERIVATION FOR THE BEC

A. Scaling Law Assumption

Throughout this section we assume that the channel W is the BEC(z) where $z \in [0, 1]$. To avoid cumbersome notation, let us define

$$p_n(z, a, b) = \Pr(Z_n \in [a, b]), \quad (26)$$

where Z_n is the Bhattacharyya process of the BEC(z). We start by noticing that by (25) the function $p_n(z, a, b)$ satisfies the following recursion

$$p_{n+1}(z, a, b) = \frac{p_n(z^2, a, b) + p_n(1 - (1 - z)^2, a, b)}{2}, \quad (27)$$

with

$$p_0(z, a, b) = \mathbb{1}_{\{z \in [a, b]\}}. \quad (28)$$

More generally, one can easily observe the following. Let $g : [0, 1] \rightarrow \mathbb{R}$ be an arbitrary bounded function. Define the functions $\{g_n\}_{n \in \mathbb{N}}$ as

$$g_n(z) = \mathbb{E}[g(Z_n)]. \quad (29)$$

Note here that in (29) the parameter z is the starting point of the process Z_n , i.e., $Z_0 = z$. The functions $\{g_n\}_{n \in \mathbb{N}}$ satisfy the following recursion for $n \in \mathbb{N}$

$$g_{n+1}(z) = \frac{g_n(z^2) + g_n(1 - (1 - z)^2)}{2}. \quad (30)$$

This observation motivates us to define the *polar operator*, call it T , as follows. Let \mathcal{B} be the space of bounded measurable functions over $[0, 1]$. The polar operator $T : \mathcal{B} \rightarrow \mathcal{B}$ maps a function $g \in \mathcal{B}$ to another function in \mathcal{B} in the following way

$$T(g) = \frac{g(z^2) + g(1 - (1 - z)^2)}{2}. \quad (31)$$

It is now clear that

$$\mathbb{E}[g(Z_n)] = T \circ T \circ \dots \circ T(g) \triangleq T^n(g). \quad (32)$$

In this new setting, our objective is to study the limiting behavior of the functions $T^n(g)$ when g is a simple function as in (28). This task is intimately related to studying the largest eigenvalues of the polar operator T and their corresponding eigenfunctions. In this regard, to keep things in a simple and manageable setting, we first consider finite-dimensional approximations of T . This is done by discretizing the unit interval into very small sub-intervals with the same length and by assuming that T operates on all the points of these sub-intervals in the same way. More concretely, consider a (large) number $L \in \mathbb{N}$ and let the numbers x_i , $i \in \{0, 1, \dots, L-1\}$ be defined as $x_i = \frac{i}{L-1}$. Hence, the unit interval $[0, 1]$ can be thought of as the union of the small sub-intervals $[x_i, x_{i+1}]$. Now, for simplicity assume that g is a (piece-wise) continuous function on $[0, 1]$. Intuitively, by assuming L to be large, we expect that the value of g is the same throughout each of the intervals $[x_i, x_{i+1}]$. Such an assumption seems also reasonable for the function $T(g)$ given in (31). We can approximate the function g as an L dimensional vector

$$g_L \approx [g(x_0), g(x_1), \dots, g(x_{L-1})]. \quad (33)$$

In this way, we expect that the function $T(g)$ can be well approximated by a matrix multiplication

$$T \approx g_L T_L, \quad (34)$$

where the $L \times L$ matrix T_L is defined as follows. Let $T_L(i, j)$ be an element of T_L in the i -th row and the j -th column. Define $T_L(1, 1) = T_L(L, L) = 1$ and for the other elements of T_L we let

$$T_L(i, j) = \begin{cases} \frac{1}{2}, & \text{if } j = \lfloor L(\frac{i}{L})^2 \rfloor, \\ \frac{1}{2}, & \text{if } j = \lceil L(1 - (1 - \frac{i}{L})^2) \rceil, \\ 0, & \text{o.w.} \end{cases} \quad (35)$$

L	1000	2000	4000	8000
$\lambda_2(L)$	0.8227	0.8240	0.8248	0.8253
$\lambda_3(L)$	0.6878	0.6958	0.7012	0.7046

TABLE I
VALUES OF $\lambda_2(L)$ AND $\lambda_3(L)$, WHICH CORRESPOND TO THE SECOND AND THIRD LARGEST EIGENVALUES OF T_L (IN ABSOLUTE VALUE), ARE COMPUTED NUMERICALLY FOR DIFFERENT VALUES OF L .

As an example, the matrix T_L for $L = 10$ has the following form

$$T_{10} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

All the rows of T_L sum up to 1. Hence, an application of the Perron-Frobenius theorem [14] shows that the eigenvalues of T_L are all inside the interval $[-1, +1]$. Also, it is easy to see that T_L has a trivial eigenvalue equal to $\lambda_0 = 1$ with two corresponding eigenvectors

$$\begin{aligned} v_0 &= (1, 0, \dots, 0), \\ v_1 &= (0, 0, \dots, 1). \end{aligned}$$

A little thought shows that the v_0 and v_1 correspond to the two extremal states of the polarization (i.e., the perfect channel and the useless channel). This can be justified by the fact that if we start from any initial vector e_p that has value one at position p and value zero elsewhere, then

$$e_p T_L^n \xrightarrow{n \rightarrow \infty} c_0 v_0 + c_1 v_1,$$

where c_0 and c_1 are positive constants. This is just a rough observation of the polarization phenomenon. In fact, by polarization one can easily guess the following. Assuming $p = zL$, we have

$$\begin{aligned} c_0 &\xrightarrow{L \rightarrow \infty} 1 - z, \\ c_1 &\xrightarrow{L \rightarrow \infty} z. \end{aligned}$$

However, we are interested in finding out how fast such a convergence is taking place. For this purpose, we look at the second and third largest eigenvalues (in absolute value) of T_L as L grows large. We denote the second largest eigenvalue of T_L by $\lambda_2(L)$ and the third largest is denoted by $\lambda_3(L)$. Table I contains the value of these eigenvalues computed numerically for several (large) values of L . It can thus be conjectured that

$$\lim_{L \rightarrow \infty} \lambda_2(L) \approx 0.826, \quad (36)$$

$$\lim_{L \rightarrow \infty} \lambda_3(L) \approx 0.705. \quad (37)$$

This belief guides us to conclude that for L growing large, if we start from any vector g which is not a multiple of the eigenvectors of T_L , then

$$gT_L^n \approx c_0 v_0 + c_1 v_1 + c_2 \lambda_2^n v_2 + O(n\lambda_3^n). \quad (38)$$

The above approximate relation indicates that for large L , the distance of gT_L^n from the limiting value is roughly equal to $c_2 \lambda_2^n$.

Now, let us go back the original polar operator T defined in (31). As we argued above, the operators T_L , for L large, are good finite-dimensional approximations of T . The (experimental) relation (38) brings us to the following assumption about T .

Assumption 1 (Scaling Assumption): There exists $\mu \in (0, \infty)$ such that, for any $z, a, b \in (0, 1)$ such that $a < b$, the limit $\lim_{n \rightarrow \infty} 2^{\frac{n}{\mu}} p_n(z, a, b)$ exists in $(0, \infty)$. We denote this limit by $p(z, a, b)$. In other words,

$$\lim_{n \rightarrow \infty} 2^{\frac{n}{\mu}} \Pr(Z_n \in [a, b]) = p(z, a, b). \quad (39)$$

We call the value μ the *scaling exponent* of polar codes for the BEC.

Note here that by (36) we expect that

$$2^{-\frac{1}{\mu}} = \lim_{L \rightarrow \infty} \lambda_2(L) \approx 0.826 \Rightarrow \frac{1}{\mu} \approx 0.275. \quad (40)$$

Let us now describe a numerical method for computing μ and $p(a, b, z)$. In this regard, we follow the approach of [11]. First we note that by (27) and the scaling law assumption we conclude that

$$2^{-\frac{1}{\mu}} p(z, a, b) = \frac{p(z^2, a, b) + p(1 - (1 - z)^2, a, b)}{2}. \quad (41)$$

Equation (41) can be solved numerically by recursion. First of all, note that the equation is invariant under multiplicative scaling of p . Also, from the equation one can naturally guess that $p(z, a, b)$ can be factorized into

$$p(z, a, b) = c(a, b)p(z), \quad (42)$$

where $p(z)$ is a solution of (41) with $p(\frac{1}{2}) = 1$. We iteratively compute μ and $p(z)$.

Initialize $p_0(z)$ –say– with $p_0(z) = 4z(1 - z)$ and compute recursively new estimates of $p_{n+1}(z)$ by first computing

$$\hat{p}_{n+1}(z) = p_n(z^2) + p_n(1 - (1 - z)^2),$$

and then by normalizing $p_{n+1}(z) = \hat{p}_{n+1}(z)/\hat{p}_{n+1}(\frac{1}{2})$, so that $p_{n+1}(\frac{1}{2}) = 1$. We have implemented the above functional recursion numerically by discretizing the z axis. Figure 4 shows the resulting numerical approximation of $p_\infty(z)$ as obtained by iterating the above procedure until $\|p_{n+1}(z) - p_n(z)\|_\infty \leq 10^{-10}$ ($\forall z \in [0, 1]$) and by using a discretization with 10^6 equi-spaced values of z . From this recursion we also get a numerical estimate of the scaling exponent μ . In particular we expect $\hat{p}_n(1/2) \rightarrow 2^{\frac{1}{\mu}}$ as $n \rightarrow \infty$. Using this method, we obtain the estimate $1/\mu \approx 0.2757$.

As mentioned above, the function $p(a, b, z)$ differs from $p(z)$ by a multiplicative constant $c(a, b)$ that is to be found by other means. In Figure 5 we plot the functions $2^{\frac{n}{\mu}} p_n(z, a, b)$ for $a = 1 - b = \frac{1}{10}$ and different values of n . We observe

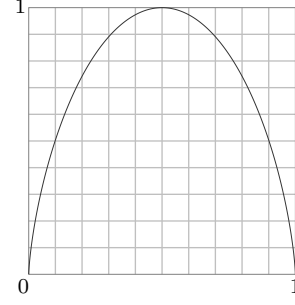


Fig. 4. The function $p(z)$ for $z \in [0, 1]$.

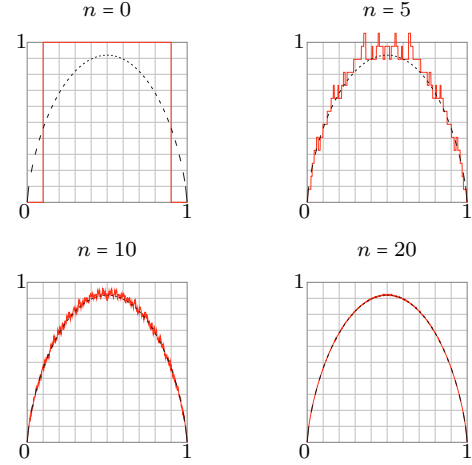


Fig. 5. The functions $2^{\frac{n}{\mu}} p_n(a, b, z)$ for various values of n . Here we have fixed $a = 1 - b = 0.9$ and $\frac{1}{\mu} = 0.2757$. In all of the four plots the dashed curve corresponds to $c(a, b)p(z)$ with $c(a, b) = 0.92$. Here, the function $p(z)$ corresponds to the numerical solution of (41).

that, as n increases these plots and the curve $c(a, b)p(z)$ with $c(a, b) = 0.92$ match very well. Even for moderate values of n (such as $n = 10$) we observe that the curves have a fairly good agreement.

Let us now see what the scaling law assumption implies about the finite-length behavior of polar codes. For simplicity, we assume that communication takes place on the BEC($\frac{1}{2}$). We are given a target error probability P_e and want to achieve a rate at least R . What block-length N should we choose?

Consider the process Z_n with $z = \frac{1}{2}$. It is easy to see that the set of possible values that Z_n takes in $[0, 1]$ is symmetric around $z = \frac{1}{2}$. Now, according to the scaling law for $x \in [0, \frac{1}{2}]$, there is a constant $p(\frac{1}{2}, x, \frac{1}{2}) \triangleq c(x)$ such that

$$\Pr(Z_n \in [x, \frac{1}{2}]) \approx c(x) 2^{-\frac{n}{\mu}}, \quad (43)$$

As a result, noticing the fact that Z_n is symmetric around the point $z = \frac{1}{2}$ we get

$$\Pr(Z_n \in [0, x]) \leq \frac{1}{2} - c(x) 2^{-\frac{n}{\mu}}. \quad (44)$$

From the construction procedure of polar codes (and specially relation (24)), we know the following. Let $z(1) \leq z(2) \leq \dots \leq z(N)$ be a re-ordering of the N possible outputs of Z_n in an

ascending order. Then, the error probability of a polar code with rate R is bounded from below by

$$P_e \geq 1 - \sqrt{1 - z(N.R)^2} \geq \frac{z(N.R)^2}{2}. \quad (45)$$

So in order to achieve error probability P_e , we should certainly have $\frac{z(N.R)^2}{2} \leq P_e$ or $z(N.R) \leq \sqrt{2P_e}$. Hence, by using (44) we deduce that

$$\begin{aligned} R &\leq \Pr(Z_n \in [0, \sqrt{2P_e}]) \\ &\leq \frac{1}{2} - c(\sqrt{2P_e})2^{-\frac{n}{\mu}} \\ &= \frac{1}{2} - c(\sqrt{2P_e})N^{-\frac{1}{\mu}}, \end{aligned}$$

and finally,

$$N \geq \left(\frac{c(\sqrt{2P_e})}{\frac{1}{2} - R} \right)^\mu. \quad (46)$$

Now, from the above calculations we know that $\frac{1}{\mu} \approx 0.2757$ as a result for the channel $W = \text{BEC}(\frac{1}{2})$ we have

$$N \geq \Theta\left(\frac{1}{(I(W) - R)^{3.627}}\right). \quad (47)$$

In the next section, we provide methods that analytically validate the above observations. We also extend some of these observations to other BMS channels.

IV. ANALYTICAL APPROACH: FROM BOUNDS FOR THE BEC TO UNIVERSAL BOUNDS FOR BMS CHANNELS

In this section we provide a rigorous basis for the observations that were derived in the previous section. Proving the full picture of Section III is beyond what we achieve here, but, we come up with close and useful bounds.

A. Characterization of μ for the BEC

We provide two approaches, that exploit different techniques, to compute the scaling exponent μ for the BEC. The first approach is based on a more careful look at equation (31). We observe that simple bounds can be derived on the largest nontrivial eigenvalue of the polar operator T by carefully analyzing the effect of T on some suitably chosen test functions. This approach provides us with a sequence of bounds on μ . We conjecture (and observe empirically) that these bounds indeed converge to the value of μ that is computed in Section III. The second approach considers different compositions of the two operations z^2 and $2z - z^2$ and analyzes the asymptotic behavior of these compositions. This approach provides us with a close lower bound on μ .

1) *First Approach:* Consider the polar operator defined in (31). The objective here is to compute the largest eigenvalues of T . Specifically, we want to find the largest solutions of

$$T(f) = \lambda f. \quad (48)$$

A check shows that both $f(z) = z$ and $f(z) = 1$ are eigenfunctions associated to the eigenvalue $\lambda = 1$. Perhaps more interestingly, let us look at the eigenvalues of T inside the interval $(0, 1)$. Intuitively, equation (41), together with the

scaling law, can be reformulated as follows. The operator T has an eigenvalue $\lambda = 2^{-\frac{1}{\mu}}$ and a corresponding eigenfunction $p(z)$ such that if we take any step function $f(z) = \mathbb{1}_{\{z \in [a, b]\}}$, then

$$\lambda^{-n} T^n(f) \xrightarrow{n \rightarrow \infty} c(a, b) p(z). \quad (49)$$

In fact, if the scaling law is true, then we naturally expect that (49) holds for a much larger class of functions rather than the class of step functions. Heuristic arguments of the previous section also suggest that (49) holds for all (piecewise) continuous functions $f(z)$ with $f(0) = f(1) = 0$.

Motivated by this picture, one approach to find bounds on the eigenvalue consists of the following two steps: (1) choose a suitable “test function” $f(z)$ for which we can provide good bounds on the behavior of $T^n(f)$ and (2) turn these bounds into bounds on the corresponding eigenvalue (or μ). With this in mind, for a generic test function $f(z) : [0, 1] \rightarrow [0, 1]$, let us define the sequence of functions $\{f_n(z)\}_{n \in \mathbb{N}}$ as $f_n : [0, 1] \rightarrow [0, 1]$ and for $z \in [0, 1]$,

$$f_n(z) \triangleq \mathbb{E}[f(Z_n)] = T^n(f). \quad (50)$$

Here, note that for $z \in [0, 1]$ the value of $f_n(z)$ is a deterministic value that is dependent on the process Z_n with the starting value $Z_0 = z$. Let us now recall once more the recursive relation of the functions f_n :

$$\begin{aligned} f_0(z) &= f(z), \\ f_n(z) &= \frac{f_{n-1}(z^2) + f_{n-1}(1 - (1 - z)^2)}{2}. \end{aligned} \quad (51)$$

In order to find lower and upper bounds on the speed of decay of the sequence f_n , we define sequences of numbers $\{a_m\}_{m \in \mathbb{N}}$ and $\{b_m\}_{m \in \mathbb{N}}$ as

$$a_m = \inf_{z \in [0, 1]} \frac{f_{m+1}(z)}{f_m(z)}, \quad (52)$$

$$b_m = \sup_{z \in [0, 1]} \frac{f_{m+1}(z)}{f_m(z)}. \quad (53)$$

Lemma 3: Fix $m \in \mathbb{N}$. For all $n \geq m$ and $z \in [0, 1]$, we have

$$(a_m)^{n-m} f_m(z) \leq f_n(z) \leq (b_m)^{n-m} f_m(z). \quad (54)$$

Furthermore, the sequence a_m is an increasing sequence and the sequence b_m is a decreasing sequence.

Proof: Here, we only prove the left-hand side of (54) and note that the right-hand side follows similarly. The proof goes by induction on $n - m$. For $n - m = 0$ the result is trivial. Assume that the relation (54) holds for a $n - m = k$, i.e., for $z \in [0, 1]$ we have

$$(a_m)^k f_m(z) \leq f_{m+k}(z). \quad (55)$$

We show that (54) is then true for $k + 1$ and $z \in [0, 1]$. We have

$$\begin{aligned} f_{m+k+1}(z) &\stackrel{(a)}{=} \frac{f_{m+k}(z^2) + f_{m+k}(1 - (1 - z)^2)}{2} \\ &\stackrel{(b)}{\geq} \frac{(a_m)^k f_m(z^2) + (a_m)^k f_m(1 - (1 - z)^2)}{2} \\ &= (a_m)^k f_{m+1}(z) \end{aligned}$$

m	0	2	4	6	10
a_m	0.75	0.7897	0.8074	0.8190	0.8239
$\log a_m$	-0.4150	-0.3406	-0.3086	-0.2880	-0.2794

TABLE II

THE VALUES OF a_m CORRESPONDING TO THE TEST FUNCTION $f_0 = z(1-z)$ ARE NUMERICALLY COMPUTED FOR SEVERAL CHOICES OF m .

$$\begin{aligned}
&= (a_m)^k \frac{f_{m+1}(z)}{f_m(z)} f_m(z) \\
&\geq (a_m)^k \left[\inf_{z \in [0,1]} \frac{f_{m+1}(z)}{f_m(z)} \right] f_m(z) \\
&= (a_m)^{k+1} f_m(z).
\end{aligned}$$

Here, (a) follows from (51) and (b) follows from the left-side inequality in (55), and hence the lemma is proved via induction. ■

Let us now begin searching for suitable test functions, i.e., candidates for $f(z)$ that provide us with good lower and upper bounds a_m and b_m . We expect that having a polynomial test function might be slightly preferable. This is due to the fact that if f is a polynomial, then $T^n(f)$ is also a polynomial and computing a_m and b_m is equivalent to finding roots of polynomials which is a manageable task. Of course the simplest polynomial that takes the value 0 on $z = 0, 1$ is $f_0(z) = z(1-z)$. Hence, let us take our test function as $f(z) = f_0(z) = z(1-z)$ and consider the corresponding sequence of functions $\{f_n(z)\}_{n \in \mathbb{N}}$,

$$f_n(z) = \mathbb{E}[Z_n(1 - Z_n)] = T^n(f_0). \quad (56)$$

A moment of thought shows that with $f_0 = z(1-z)$ the function $2^n f_n$ is a polynomial of degree 2^{n+1} with integer coefficients. Let us first focus on computing the value of a_m for $m \in \mathbb{N}$. If the relation (49) holds true, then we expect that the value of a_m converges to $\lambda = 2^{-\frac{1}{\mu}}$ as m grows large.

Remark 4: One can compute the value of a_m by finding the extreme points of the function $\frac{f_{m+1}}{f_m}$ (i.e., finding the roots of the polynomial $g_m = f'_{m+1}f_m - f_{m+1}f'_m$) and checking which one gives the global minimum. Assuming $f_0 = z(1-z)$, for small values e.g., $m = 0, 1$, pen and paper suffice. For higher values of m , we can automatize the process: all these polynomials have rational coefficients and therefore it is possible to determine the number of real roots exactly and to determine their value to any desired precision. This task can be accomplished precisely by computing so-called Sturm chains (see Sturm's Theorem [17]). Computing Sturm chains is equivalent to running Euclid's algorithm starting with the second and third derivative of the original polynomial. Hence, we can find the value of a_m analytically to any desired precision. Table II contains the numerical value of a_m up to precision 10^{-4} for $m \leq 10$. As the table shows, the values a_m are increasing (see Lemma 3), and we conjecture that they converge to $2^{-0.2757} = 0.8260$, the corresponding value for the channel BEC.

Let us now focus on computing the value of b_m . On the negative side, for the specific test function $f(z) = z(1-z)$ we obtain $b_m = 1$ for $m \in \mathbb{N}$ and therefore the upper bounds of (53) are of trivial use. In fact, it is not hard to show that

m	0	2	4	6	8
b_m	0.8312	0.8294	0.8279	0.8268	0.8264
$\log b_m$	-0.2663	-0.2699	-0.2725	-0.2744	-0.2751

TABLE III

THE VALUES OF b_m CORRESPONDING TO $f_0 = (z(1-z))^{\frac{2}{3}}$ ARE NUMERICALLY COMPUTED FOR SEVERAL CHOICES OF m .

if we plug in any polynomial as the test function then we get $b_m = 1$ for any m . On the positive side, we can consider other test functions that result in non-trivial values for b_m . The problem with non-polynomial functions is that methods such as the Sturm-chain method no longer apply here. Hence, finding the precise value of b_m up to a desired precision can be a difficult task and we lose the analytical tractability of b_m . As an example, choose

$$f_0(z) = z^\alpha(1-z)^\beta, \quad (57)$$

for some choice of $\alpha, \beta \in (0, 1)$. Then, from (53) we have

$$b_0 = \sup_{z \in [0,1]} \frac{f_1(z)}{f_0(z)} = \sup_{z \in [0,1]} \frac{z^\alpha(1+z)^\beta + (2-z)^\alpha(1-z)^\beta}{2}. \quad (58)$$

By letting $\alpha = \beta = \frac{2}{3}$, we numerically get $b_0 = 0.8312$ which is already a close bound for λ . This suggests that the test function $f_0(z) = f(z) = (z(1-z))^{\frac{2}{3}}$ is suitable candidate for obtaining good upper bounds b_m . For this specific test function, the value of b_m for various values of m has been numerically computed in Table III. As we observe from Table III, even for moderate values of m the (numerical) bound b_m is very close to the true "value" of λ .

Finally, let us relate the bounds a_m and b_m to bounds on the functions $p_n(a, b, z)$. We have

Lemma 5: Consider the test function $f(z) = z(1-z)$ and the corresponding sequence of function f_n defined in (51). Let $a, b \in (0, 1)$ be such that $\sqrt{a} \leq 1 - \sqrt{1-b}$. Then, there are constants $c_1, c_2 > 0$ such that for any $z \in (0, 1)$

$$\begin{aligned}
\frac{1}{n} \log f_n(z) - \frac{c_1 \log n}{n} &\leq \frac{1}{n} \log \Pr(Z_n \in [a, b]) \\
&\leq \frac{1}{n} \log f_n(z) + \frac{c_2}{n}. \quad (59)
\end{aligned}$$

Also, for the test function $f(z) = (z(1-z))^{\frac{2}{3}}$ and the corresponding sequence f_n , defined in (51), we have for $a, b \in (0, 1)$

$$\frac{1}{n} \log \Pr(Z_n \in [a, b]) \leq \frac{1}{n} \log f_n(z) + \frac{c_3}{n}, \quad (60)$$

where c_3 is a positive constant.

We can now easily conclude that

Corollary 6: Fix $m \in \mathbb{N}$. For $a, b \in [0, 1]$ such that $\sqrt{a} \leq 1 - \sqrt{1-b}$ and $n \leq m$ we have

$$\log a_m + O\left(\frac{\log n}{n}\right) \leq \frac{1}{n} \log \Pr(Z_n \in [a, b]) \leq \log b_m + O\left(\frac{1}{n}\right), \quad (61)$$

where a_m is defined in (52) with the test function $f(z) = z(1-z)$ (see Table II), and b_m is defined in (108) with the test function $f(z) = (z(1-z))^{\frac{2}{3}}$ (see Table III).

Remark 7: We expect that the result of Lemma 5 holds for any choice of a and b such that $a < b$. That is, the condition $\sqrt{a} \leq 1 - \sqrt{1-b}$ is not a serious condition and is just given to ease out the proof.

2) *Second Approach:* Throughout this section we will prove the following theorem.

Theorem 8: We have

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \left\{ \int_0^1 \Pr(Z_n \in [a, b]) dz \right\} \geq \frac{1}{2 \ln 2} - 1 \approx -0.2787. \quad (62)$$

Let us now explain, at the intuitive level, the main consequence of Theorem 8. By using the scaling law assumption, and specifically (38) and (39), we have that $\int_0^1 \Pr(Z_n \in [a, b]) dz \approx \int_0^1 2^{-\frac{n}{2}} p(z, a, b) dz + o(2^{-\frac{n}{2}})$. This relation together with (62) results that $\mu \geq \frac{1}{2 \ln 2} - 1 \approx -0.2787$. For the sake of brevity, we do not address here further (analytic) conclusions of Theorem 8 and we refer the reader to [12].

To proceed with the proof, let us recall from Section I-D the definition of Z_n (for the BEC) in terms of the sequence $\{B_n\}_{n \in \mathbb{N}}$. We start by $Z_0 = z$ and

$$Z_{n+1} = \begin{cases} Z_{n-1}^2 & \text{if } B_n = 1, \\ 2Z_{n-1} - Z_{n-1}^2 & \text{if } B_n = 0. \end{cases} \quad (63)$$

Hence, by considering the two maps $t_0, t_1 : [0, 1] \rightarrow [0, 1]$ defined as

$$t_0(z) = 2z - z^2, t_1(z) = z^2, \quad (64)$$

the value of Z_n is obtained by applying t_{B_n} on the value of Z_{n-1} , i.e.,

$$Z_n = t_{B_n}(Z_{n-1}). \quad (65)$$

The same rule applies for obtaining the value of Z_{n-1} from Z_{n-2} and so on. Thinking this through recursively, the value of Z_n is obtained from the starting point of the process, $Z_0 = z$, via the following (random) maps.

Definition 9: For each $n \in \mathbb{N}$ and a realization $(b_1, \dots, b_n) \triangleq \omega_n \in \Omega_n$ define the map ϕ_{ω_n} by

$$\phi_{\omega_n} = t_{b_n} \circ t_{b_{n-1}} \circ \dots \circ t_{b_1}. \quad (66)$$

Also, let Φ_n be the set of all such n -step maps.

As a result, an equivalent description of the process Z_n is as follows. At time n the value of Z_n is obtained by picking uniformly at random one of the functions in $\phi_{\omega_n} \in \Phi_n$ and assigning the value $\phi_{\omega_n}(z)$ to Z_n . Consequently we have,

$$\Pr(Z_n \in [a, b]) = \sum_{\phi_{\omega_n} \in \Phi_n} \frac{1}{2^n} \mathbb{1}_{\{\phi_{\omega_n}(z) \in [a, b]\}}. \quad (67)$$

Using (67), it is apparent that in order to analyze the behavior of the quantity $\frac{1}{n} \log \Pr(Z_n \in [a, b])$ as n grows large, it is necessary to characterize the asymptotic behavior of the random maps ϕ_{ω_n} . Continuing the theme of Definition 9, we can assign to each realization of the infinite sequence $\{B_k\}_{k \in \mathbb{N}}$, denoted by $\{b_k\}_{k \in \mathbb{N}}$, a sequence of maps $\phi_{\omega_1}(z), \phi_{\omega_2}(z), \dots$, where $\omega_i \triangleq (b_1, \dots, b_i)$. We call the sequence $\{\phi_{\omega_k}\}_{k \in \mathbb{N}}$ the corresponding sequence of maps for the realization $\{b_k\}_{k \in \mathbb{N}}$. We also use the realization $\{b_k\}_{k \in \mathbb{N}}$ and its corresponding $\{\phi_{\omega_k}\}_{k \in \mathbb{N}}$ interchangeably. Let us now focus on the asymptotic characteristics of the functions ϕ_{ω_n} . Firstly,

since $\{\phi_{\omega_n}(z)\}_{\omega_n \in \Omega_n}$ has the same law as Z_n starting at z , we conclude that for $z \in [0, 1]$ with probability one, the quantity $\lim_{k \rightarrow \infty} \phi_{\omega_k}(z)$ takes on a value in the set $\{0, 1\}$. In Figure 6 the functions ϕ_{ω_n} are plotted for a random realization. As it is apparent from the figure, the functions ϕ_{ω_n} seem to converge point-wise to a jump function (i.e., a sharp rise from 0 to 1). As intuitive justification of this fact is as follows. Consider a random function ϕ_{ω_n} . Due to polarization, as n grows large, almost all the values that this function takes are very close to 0 or 1. This function is also increasing and continuous (more precisely, it is a polynomial). A little thought reveals that the only choice to imagine for ϕ_{ω_n} is a very sharp rise from being almost 0 to almost 1. The formal and complete statement is given as follows..

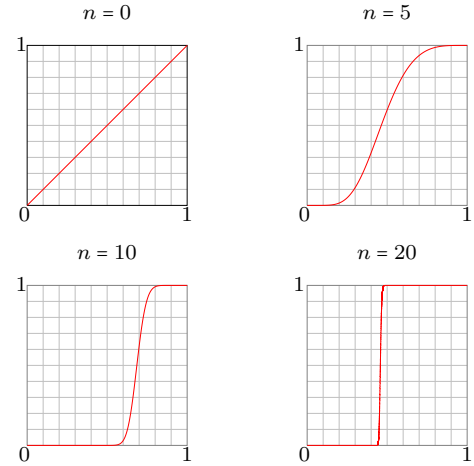


Fig. 6. The functions ϕ_{ω_n} associated to a random realization are plotted. As we see as n grows large, the functions ϕ_{ω_n} converge point-wise to a step function.

Lemma 10 (Almost every realization has a threshold point): For almost every realization of $\omega \triangleq \{b_k\}_{k \in \mathbb{N}} \in \Omega$, there exists a point $z_\omega^* \in [0, 1]$, such that

$$\lim_{n \rightarrow \infty} \phi_{\omega_n}(z) \rightarrow \begin{cases} 0 & z \in [0, z_\omega^*) \\ 1 & z \in (z_\omega^*, 1] \end{cases}$$

Furthermore, z_ω^* has uniform distribution on $[0, 1]$. We call the point z_ω^* the threshold point of the realization $\{b_k\}_{k \in \mathbb{N}}$ or the threshold point of its corresponding sequence of maps $\{\phi_{\omega_k}\}_{k \in \mathbb{N}}$.

Looking more closely at (67), by the above lemma we conclude that as n grows large, the maps ϕ_{ω_n} that activate the identity function $\mathbb{1}_{\{\cdot\}}$ must have their threshold point sufficiently close to z . Let us now give an intuitive discussion about the idea behind the proof of Theorem 8. By using (67) we can write

$$\begin{aligned} \Pr(Z_n \in [a, b]) &= \sum_{\phi_{\omega_n} \in \Phi_n} \frac{1}{2^n} \mathbb{1}_{\{\phi_{\omega_n}(z) \in [a, b]\}} \\ &= \sum_{\phi_{\omega_n} \in \Phi_n} \frac{1}{2^n} \mathbb{1}_{\{z \in [\phi_{\omega_n}^{-1}(a), \phi_{\omega_n}^{-1}(b)]\}}. \end{aligned} \quad (68)$$

Hence by Lemma 10, for a large choice of n the intervals $[\phi_{\omega_n}^{-1}(a), \phi_{\omega_n}^{-1}(b)]$ have a very short length and are distributed almost uniformly along $[0, 1]$. Now, if we assume that the length of the intervals $[\phi_{\omega_n}^{-1}(a), \phi_{\omega_n}^{-1}(b)]$ is very close to their average, then we can replace the average in (68) by the average length of $[\phi_{\omega_n}^{-1}(a), \phi_{\omega_n}^{-1}(b)]$. That is,

$$\Pr(Z_n \in [a, b]) \approx \mathbb{E}[\phi_{\omega_n}^{-1}(b) - \phi_{\omega_n}^{-1}(a)].$$

So intuitively, all that remains is to compute the average length of the random intervals $[\phi_{\omega_n}^{-1}(a), \phi_{\omega_n}^{-1}(b)]$.

In fact we are not able to make all these heuristics precise for the point-wise values $\frac{1}{n} \log \Pr(Z_n \in [a, b])$. Nonetheless, the picture is naturally precise for the average of $\Pr(Z_n \in [a, b])$ over $z \in [0, 1]$, i.e.,

$$\frac{1}{n} \log \left\{ \int_0^1 \Pr(Z_n \in [a, b]) dz \right\}. \quad (69)$$

To see this, we proceed as follows. By (68) we have

$$\begin{aligned} \int_0^1 \Pr(Z_n \in [a, b]) dz &= \int_0^1 \left\{ \sum_{\phi_{\omega_n}} \frac{1}{2^n} \mathbb{1}_{\{z \in \phi_{\omega_n}^{-1}[a, b]\}} \right\} dz \\ &= \sum_{\phi_{\omega_n}} \frac{1}{2^n} \left\{ \int_0^1 \mathbb{1}_{\{z \in \phi_{\omega_n}^{-1}[a, b]\}} dz \right\} \\ &= \mathbb{E}[\phi_{\omega_n}^{-1}(b) - \phi_{\omega_n}^{-1}(a)], \end{aligned}$$

and by applying $\frac{1}{n} \log(\cdot)$ to both sides we have

$$\begin{aligned} \frac{1}{n} \log \left\{ \int_0^1 \Pr(Z_n^z \in [a, b]) dz \right\} &= \frac{1}{n} \log \mathbb{E}[\phi_{\omega_n}^{-1}(b) - \phi_{\omega_n}^{-1}(a)] \\ &\geq \frac{1}{n} \mathbb{E}[\log(\phi_{\omega_n}^{-1}(b) - \phi_{\omega_n}^{-1}(a))], \end{aligned} \quad (70)$$

where in the last step we have used Jensen's inequality. The value of $\lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{E}[\log(\phi_{\omega_n}^{-1}(b) - \phi_{\omega_n}^{-1}(a))]$ can be computed precisely.

Lemma 11: We have

$$\lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{E}[\log(\phi_{\omega_n}^{-1}(b) - \phi_{\omega_n}^{-1}(a))] = \frac{1}{2 \ln 2} - 1 \approx -0.2787.$$

As a result, we have

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \left\{ \int_0^1 \Pr(Z_n \in [a, b]) dz \right\} \geq \frac{1}{2 \ln 2} - 1.$$

The result of Theorem 8 provides a lower bound that is very close to the value we obtained in Section III but is not exactly equal. This is because we have used Jensen's inequality in (70).

B. Speed of Polarization for General BMS Channels

For a BMS channel W , there is no simple 1-dimensional recursion for the process Z_n as for the BEC. However, by using (12) and (13), we can provide bounds on how Z_n evolves:

$$Z_{n+1} \begin{cases} = Z_n^2 & ; \text{if } B_n = 1, \\ \in [Z_n \sqrt{2 - Z_n^2}, 2Z_n - Z_n^2] & ; \text{if } B_n = 0. \end{cases} \quad (71)$$

As a warm-up, we notice that similar techniques as used in Section IV-A1 are applicable to provide general lower and upper bounds. For instance, to find upper bounds we can proceed

as follows. For any non-negative function $g : [0, 1] \rightarrow \mathbb{R}^+$ such that $g(0) = g(1) = 0$ let

$$L_g = \sup_{z \in (0, 1), y \in [z\sqrt{2-z^2}, z(2-z)]} \frac{g(z^2) + g(y)}{2g(z)}.$$

Similar to the discussion in Section IV-A1, we can show that for the process $Z_n = Z(W_n)$ we have

$$\mathbb{E}[g(Z_n)] \leq c L_g^n, \quad (72)$$

where $c = \sup_{z \in [0, 1]} g(z)$ is a constant. Hence, using the Markov inequality we have for $a, b \in (0, 1)$,

$$\frac{1}{n} \log \Pr(Z_n \in [a, b]) \leq \log L_g + O\left(\frac{1}{n}\right).$$

For example, assuming $g(z) = (z(1-z))^{\frac{2}{3}}$ we numerically obtain that $\log L_g = -0.169$. That is

$$\mathbb{E}[(Z_n(1-Z_n))^{\frac{2}{3}}] \leq 2^{-0.169n}, \quad (73)$$

and for $a, b \in (0, 1)$ we have

$$\frac{1}{n} \log \Pr(Z_n \in [a, b]) \leq -0.169 + O\left(\frac{1}{n}\right).$$

The relations of type (73) are upper bounds on the speed of polarization that hold *universally* over all the BMS channels. Let us now compute universal lower bounds. In the rest of this section, it is more convenient for us to consider another stochastic process related to W_n , which is the process⁴ $H_n = H(W_n)$. The main reason to consider H_n rather than Z_n is that the process H_n is a martingale and this martingale property will help us to use the functions $\{f_n\}_{n \in \mathbb{N}}$ defined in (51) (with the starting function $f(z) = z(1-z)$) to provide universal lower bounds on the quantity $\mathbb{E}[H_n(1-H_n)]$. We begin by introducing one further technical condition given as follows.

Definition 12: We call an integer $m \in \mathbb{N}$ *suitable* if the function $f_m(z)$, defined in (51) (with the starting function $f(z) = z(1-z)$), is concave on $[0, 1]$.

Remark 13: For small values of m , i.e., $m \leq 2$, it is easy to verify by hand that the function f_m is concave. As discussed previously, for larger values of m we can use Sturm's theorem [17] and a computer algebra system to verify this. Note that the polynomials $2^m f_m$ have integer coefficients. Hence, all the required computations can be done exactly. We have checked up to $m = 8$ that f_m is concave and we conjecture that in fact this is true for all $m \in \mathbb{N}$.

We now show that for any BMS channel W , the value of a_m , defined in (52), is a lower bound on the speed of decay of H_n provided that m is a suitable integer.

Lemma 14: Let $m \in \mathbb{N}$ be a suitable integer and W a BMS channel. We have for $n \geq m$

$$\mathbb{E}[H_n(1-H_n)] \geq (a_m)^{n-m} f_m(H(W)), \quad (74)$$

where a_m is given in (52).

Proof: We use induction on $n - m$: for $n - m = 0$ there is nothing to prove. Assume that the result of the lemma is

⁴For the BEC the processes H_n and Z_n are identical.

correct for $n - m = k$. Hence, for any BMS channel W with $H_n = H(W_n)$ we have

$$\mathbb{E}[H_{m+k}(1 - H_{m+k})] \geq (a_m)^k f_m(H(W)). \quad (75)$$

We now prove the lemma for $m - n = k + 1$. For the BMS channel W , let us recall from Section I-B that the transform $W \rightarrow (W^0, W^1)$ yields two channels W^0 and W^1 such that (11) holds. Define the process $\{(W^0)_n, n \in \mathbb{N}\}$ as the channel process that starts with W^0 and evolves as in (16). We define $\{(W^1)_n, n \in \mathbb{N}\}$ similarly. Let us also define the two processes $H_n^0 = H((W^0)_n)$ and $H_n^1 = H((W^1)_n)$. We have,

$$\begin{aligned} & \mathbb{E}[H_{m+k+1}(1 - H_{m+k+1})] \\ & \stackrel{(a)}{=} \frac{\mathbb{E}[H_{m+k}^0(1 - H_{m+k}^0)] + \mathbb{E}[H_{m+k}^1(1 - H_{m+k}^1)]}{2} \\ & \stackrel{(b)}{\geq} (a_m)^k \frac{f_m(H(W^0)) + f_m(H(W^1))}{2} \\ & \stackrel{(c)}{\geq} (a_m)^k \frac{f_m(1 - (1 - H(W))^2) + f_m(H(W)^2)}{2} \\ & \stackrel{(d)}{=} (a_m)^k f_{m+1}(H(W)) \\ & = (a_m)^k \frac{f_{m+1}(H(W))}{f_m(H(W))} f_m(H(W)) \\ & \geq (a_m)^k \left[\inf_{h \in [0,1]} \frac{f_{m+1}(h)}{f_m(h)} \right] f_m(H(W)) \\ & \stackrel{(e)}{=} (a_m)^{m+1} f_m(H(W)). \end{aligned}$$

In the above chain of inequalities, relation (a) follows from the fact that W_m has 2^m possible outputs among which half of them are branched out from W^0 and the other half are branched out from W^1 . Relation (b) follows from the induction hypothesis given in (75). Relation (c) follows from (22), (23) and the fact that the function f_m is concave. More precisely, because f_m is concave on $[0, 1]$, we have the following inequality for any sequence of numbers $0 \leq x' \leq x \leq y \leq y' \leq 1$ that satisfy $\frac{x+y}{2} = \frac{x'+y'}{2}$:

$$\frac{f_m(x') + f_m(y')}{2} \leq \frac{f_m(x) + f_m(y)}{2}. \quad (76)$$

In particular, we set $x' = H(W)^2$, $x = H(W^1)$, $y = H(W^0)$, $y' = 1 - (1 - H(W))^2$ and we know from (22) and (23) that $0 \leq x' \leq x \leq y \leq y' \leq 1$. Hence, by (76) we obtain (c). Relation (d) follows from the recursive definition of f_m given in (51). Finally, relation (e) follows from the definition of a_m given in (52). ■

Finally in the following two parts, we rigorously relate the results obtained in previous sections to finite-length performance of polar codes. In other words, answering Question 4 is the main focus for the remaining parts of this section.

C. Universal Bounds on the Scaling Behavior of Polar Codes

1) *Universal Lower Bounds:* Consider a BMS channel W and let us assume that a polar code with block-error probability at most a given value $P_e > 0$, is required. One way to accomplish this is to ensure that the right side of (24) is less than P_e . However, this is only a sufficient condition that might not be necessary. Hence, we call the right side of

(24) *the strong reliability condition*. Numerical and analytical investigations (see [11] and [18]) suggest that once the sum of individual errors in the right side of (24) is less than 1, then it provides a fairly good estimate of P_e . In fact, the smaller the sum is the closer it is to P_e . Hence, the sum of individual errors can be considered as a fairly accurate proxy for P_e . Based on this measure of the block-error probability, we provide bounds on how the rate R scales in terms of the block-length N .

Theorem 15: For any BMS channel W with capacity $I(W) \in (0, 1)$, there exist constants $P_e, \alpha > 0$, that depend only on $I(W)$, such that

$$\sum_{i \in \mathcal{L}_{N,R}} E(W_N^{(i)}) \leq P_e, \quad (77)$$

implies

$$R < I(W) - \frac{\alpha}{N^{\frac{1}{\mu}}}, \quad (78)$$

where μ is a universal parameter lower bounded by 3.553.

Here, a few comments are in order:

(i) As we have seen above, we can obtain an increasing sequence of lower bounds, call this sequence $\{\mu_m\}_{m \in \mathbb{N}}$, for the universal parameter μ . For each m , in order to show the validity of the lower bound, we need to verify the concavity of a certain polynomial (defined in (51)) in $[0, 1]$. We explained in Remark 13 how we can accomplish this using the Sturm chain method. The lower bound for μ stated in Theorem 15 is the one corresponding to $m = 8$, an arbitrary choice. If we increase m , we get e.g., $\mu_{16} = 3.614$. We conjecture that the sequence μ_m converges to $\mu = 3.627$, the parameter for the BEC. If such a conjecture holds, then the channel BEC polarizes the fastest among the BMS channels (see Question 2).

(ii) Let P_e, α, μ be as in Theorem 15. If we require the block-error probability to be less than P_e (in the sense that the condition (77) is fulfilled), then the block-length N should be at least

$$N > \left(\frac{\alpha}{I(W) - R} \right)^\mu. \quad (79)$$

(iii) From (1) we know that the value of μ for the random linear ensemble is $\mu = 2$, which is the optimal value since the variations of the channel itself require $\mu \geq 2$. Thus, given a rate R , reliable transmission by polar codes requires a larger block-length than the optimal value.

Proof of Theorem 15: To fit the bounds of Section IV-A1 into the framework of Theorem 15, let us first introduce the sequence $\{\mu_m\}_{m \in \mathbb{N}}$ as

$$\mu_m = -\frac{1}{\log a_m}, \quad (80)$$

where a_m is defined in (52) with starting function $f(z) = z(1 - z)$. In the previous section, we have proved that for a suitable m , the speed with which the quantity $\mathbb{E}[H_n(1 - H_n)]$ decays is lower bounded by $a_m = 2^{-\frac{1}{\mu_m}}$, i.e. for $n \geq m$ we have $\mathbb{E}[H_n(1 - H_n)] \geq 2^{-\frac{(n-m)}{\mu_m}} f_m(H(W))$. To relate the strong reliability condition in (77) to the rate bound in (78), we need the following lemma.

Lemma 16: Consider a BMS channel W and assume that there exist positive real numbers γ, θ and $m \in \mathbb{N}$ such that

$\mathbb{E}[H_n(1 - H_n)] \geq \gamma 2^{-n\theta}$ for $n \geq m$. Let $\alpha, \beta \geq 0$ be such that $2\alpha + \beta = \gamma$, we have for $n \geq m$

$$\Pr(H_n \leq \alpha 2^{-n\theta}) \leq I(W) - \beta 2^{-n\theta}. \quad (81)$$

Proof: The proof is by contradiction. Let us assume the contrary, i.e., we assume there exists $n \geq m$ s.t.,

$$\Pr(H_n \leq \alpha 2^{-n\theta}) > I(W) - \beta 2^{-n\theta}. \quad (82)$$

In the following, we show that with such an assumption we reach to a contradiction. We have

$$\begin{aligned} \mathbb{E}[H_n(1 - H_n)] &= \mathbb{E}[H_n(1 - H_n) | H_n \leq \alpha 2^{-n\theta}] \Pr(H_n \leq \alpha 2^{-n\theta}) \\ &\quad + \mathbb{E}[H_n(1 - H_n) | H_n > \alpha 2^{-n\theta}] \Pr(H_n > \alpha 2^{-n\theta}). \end{aligned} \quad (83)$$

It is now easy to see that

$$\mathbb{E}[H_n(1 - H_n) | H_n \leq \alpha 2^{-n\theta}] \leq \alpha 2^{-n\theta},$$

and since $\mathbb{E}[H_n(1 - H_n)] \geq \gamma 2^{-n\theta}$, by using (83) we get

$$\mathbb{E}[H_n(1 - H_n) | H_n > \alpha 2^{-n\theta}] \Pr(H_n > \alpha 2^{-n\theta}) \geq 2^{-n\theta}(\gamma - \alpha). \quad (84)$$

We can further write

$$\begin{aligned} \mathbb{E}[(1 - H_n)] &= \mathbb{E}[1 - H_n | H_n \leq \alpha 2^{-n\theta}] \Pr(H_n \leq \alpha 2^{-n\theta}) \\ &\quad + \mathbb{E}[1 - H_n | H_n > \alpha 2^{-n\theta}] \Pr(H_n > \alpha 2^{-n\theta}), \end{aligned} \quad (85)$$

and noticing fact that $H_n \geq H_n(1 - H_n)$ we can plug (84) in (85) to obtain

$$\begin{aligned} \mathbb{E}[(1 - H_n)] &\geq \mathbb{E}[1 - H_n | H_n \leq \alpha 2^{-n\theta}] \Pr(H_n \leq \alpha 2^{-n\theta}) \\ &\quad + 2^{-n\theta}(\gamma - \alpha). \end{aligned} \quad (86)$$

We now continue by using (82) in (86) to obtain

$$\begin{aligned} \mathbb{E}[(1 - H_n)] &> (I(W) - \beta 2^{-n\theta})(1 - \alpha 2^{-n\theta}) + 2^{-n\theta}(\gamma - \alpha) \\ &\geq I(W) + 2^{-n\theta}(\gamma - \alpha(1 + I(W)) - \beta), \end{aligned}$$

and since $2\alpha + \beta = \gamma$, we get $\mathbb{E}[1 - H_n] > I(W)$. This is a contradiction since H_n is a martingale and $\mathbb{E}[1 - H_n] = I(W)$. ■

Let us now use the result of Lemma 16 to conclude the proof of Theorem 15. By Lemma 14, we have for $n \geq m$

$$\mathbb{E}[H_n(1 - H_n)] \geq 2^{-\frac{(n-m)}{\mu_m}} f_m(H(W)).$$

Thus, if we now let $\gamma = 2^{\frac{m}{\mu_m}} f_m(H(W))$ and $2\alpha = \beta = \frac{\gamma}{2}$, then by using Lemma 16 we obtain

$$\Pr(H_n \leq \frac{\gamma}{4} 2^{-\frac{n}{\mu_m}}) \leq I(W) - \frac{\gamma}{2} 2^{-\frac{n}{\mu_m}}. \quad (87)$$

Assume that we desire to achieve a rate R equal to

$$R = I(W) - \frac{\gamma}{4} 2^{-\frac{n}{\mu_m}}. \quad (88)$$

Let $\mathcal{I}_{N,R}$ be the set of indices chosen for such a rate R , i.e., $\mathcal{I}_{N,R}$ includes the $2^n R$ indices of the sub-channels with the least value of error probability. Define the set A as

$$A = \{i \in \mathcal{I}_{N,R} : H(W_N^{(i)}) \geq \frac{\gamma}{4} 2^{-\frac{n}{\mu_m}}\}. \quad (89)$$

In this regard, note that (87) and (88) imply that $|A| \geq \frac{\gamma}{4} 2^{n(1-\frac{1}{\mu_m})}$. As a result, by using (5) and (6) we obtain

$$\begin{aligned} \sum_{i \in \mathcal{I}_{N,R}} E(W_N^{(i)}) &\geq \sum_{i \in A} E(W_N^{(i)}) \geq \frac{\gamma^2}{16} 2^{n(1-\frac{1}{\mu_m})} h_2^{-1}(2^{-\frac{n}{\mu_m}}) \\ &\geq \frac{\gamma^2}{16} \frac{2^{n(1-2\frac{1}{\mu_m})}}{8n \frac{1}{\mu_m}}, \end{aligned} \quad (90)$$

where the last step follows from the fact that for $x \in [0, \frac{1}{\sqrt{2}}]$, we have $h_2^{-1}(x) \geq \frac{x}{8 \log(\frac{1}{x})}$. Thus, having a block-length $N = 2^n$, in order to have error probability (measured by (24)) less than $\frac{\gamma^2}{16} \frac{2^{n(1-2\frac{1}{\mu_m})}}{8n \frac{1}{\mu_m}}$, the rate can be at most $I(W) - \frac{\gamma}{4} 2^{-\frac{n}{\mu_m}}$.

Finally, if we let $m = 8$ (by the discussion in Remark 13, we know that $m = 8$ is suitable), then $\mu_8 = \frac{1}{-\log(a_8)} = 3.553$ and choosing

$$P_e = \inf_{n \in \mathbb{N}} \left[\sum_{i \in \mathcal{I}_{N,R}} E(W_N^{(i)}) \right], \quad (91)$$

where R is given in (88), then it is easy to see from (90) that $P_e > 0$ (since $\frac{1}{\mu_8} < \frac{1}{2}$) and furthermore, to have block-error probability less than P_e the rate should be less than R given in (88).

2) *Universal Upper Bounds:* In this part, we provide upper bounds on the required block-length of Question 4. Again, the key observation here is the upper-bounds on the speed of polarization, e.g. the bounds derived in Table III for the BEC and the universal bound (73).

Theorem 17: Let $Z_n = Z(W_n)$ be the Bhattacharyya process associated to a BMS channel W . Assume that for $n \in \mathbb{N}$ we have

$$\mathbb{E}[(Z_n(1 - Z_n))^\alpha] \leq \beta 2^{-\rho n}, \quad (92)$$

where α, β, ρ are positive constants and $\alpha < 1$. Then, the block-length N required to achieve an error probability $P_e > 0$ at a given rate $R < I(W)$ is bounded above by

$$\begin{aligned} \log N &\leq (1 + \frac{1}{\rho}) \log \frac{1}{d} + \\ &c_4 (\log(\log \frac{3}{d}))^2 + c_5 \log(\log(\frac{2}{P_e})) \log(\log \frac{3}{d}), \end{aligned} \quad (93)$$

where $d = I(W) - R$ and c_4, c_5 are positive constants that depend on α, β, ρ .

Before proceeding with the proof of Theorem 17, let us note a few comments:

(i) In the previous sections we have computed several candidates for the value ρ required in Theorem 17. As an example, using the universal candidate for ρ obtained in (73) (i.e., $\rho = 0.169$), we obtain the following corollary.

Corollary 18: For any BMS channel W , the block-length N required to achieve a rate $R < I(W)$ scales at most as

$$N \leq \Theta\left(\frac{1}{(I(W) - R)^7}\right). \quad (94)$$

One important consequence of this corollary is that polar codes require a block-length that scales polynomially in terms of the gap to capacity⁵.

⁵The fact that polar codes need a polynomial block-length in terms of the gap to capacity is also proven in the recent independently-derived result of [19].

(ii) As we will see in the proof of Theorem 17, the result of this theorem is also valid if we replace P_e with the sum of Bhattacharyya values of the channels that correspond to the good indices (this sum is indeed an upper bound for P_e).

Proof of Theorem 17: Throughout the proof we will be using two key lemmas (Lemma 19 and Lemma 20) that are stated in the appendices. Let

$$d = I(W) - R. \quad (95)$$

We define $n_0 \in \mathbb{N}$ to be

$$n_0 = \left\lceil \frac{1}{\rho} \log \frac{3(1+c_1)(1+2c_2c_3)}{d} \right\rceil, \quad (96)$$

where the constants c_1 , c_2 and c_3 are given in Lemmas 19, 20 and 21, respectively. As a result of Lemma 19 and (96), we have for $n \geq n_0$

$$\Pr(Z_n \leq \frac{1}{2}) \geq R + \frac{2}{3}d. \quad (97)$$

We also define the set \mathcal{A} as follows. Let $N_0 = 2^{n_0}$ and

$$\mathcal{A} = \{i \in \{0, \dots, N_0 - 1\} : Z(W_{N_0}^{(i)}) \leq \frac{1}{2}\}. \quad (98)$$

In other words \mathcal{A} is the set of indices at level n_0 of the corresponding infinite binary tree of W (see Section I-C) whose Bhattacharyya parameter is not so large. Also, from (97) the set \mathcal{A} contains more than a fraction R of all the sub-channels at level n_0 . The idea is then to go further down through the infinite binary tree at a level $n_0 + n_1$ (the value of n_1 will be specified shortly). We then observe that the sub-channels at level $n_0 + n_1$ that are branched out from the set \mathcal{A} are polarized to a great extent in the sense that sum of their Bhattacharyya parameters is below P_e (see Figure 7 for a schematic illustration of the idea).

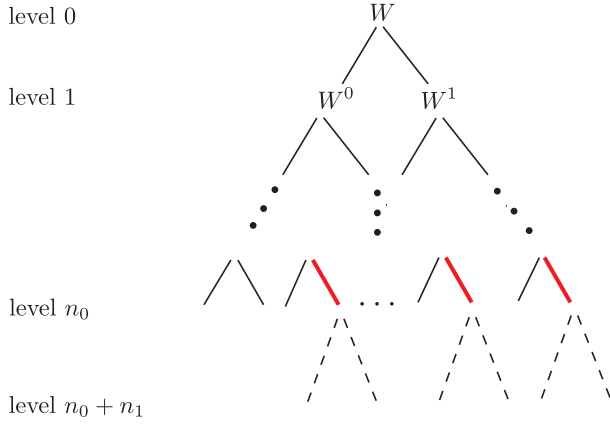


Fig. 7. The infinite binary tree of channel W . The edges that are colored red at level n_0 of this tree correspond to the sub-channels at level n_0 whose Bhattacharyya parameter is less than $\frac{1}{2}$ (i.e., the set \mathcal{A}). The idea is then to focus on these “red” indices. We consider the sub-channels that are branched out from these red indices at a level $n_0 + n_1$ (as shown in the figure). By a careful choice of n_1 , we observe that these specific sub-channels at level $n_0 + n_1$ are greatly polarized in the sense that sum of their Bhattacharyya parameters is less than P_e . We also show that the fraction of these sub-channels is larger than R .

We proceed by finding a suitable candidate for n_1 . Our objective is to choose n_1 large enough s.t. there is a set of

indices at level $n_0 + n_1$ with the following properties: (i) sum of the Bhattacharyya parameters of the sub-channels in this set is less than P_e and (ii) the cardinality of this set is at least $R2^{n_0+n_1}$. In what follows, we will first use the hypothesis of Lemma 20 to give a candidate for n_1 and then we make it clear that such a candidate is suitable for our needs. Let $\{B_m\}_{m \in \mathbb{N}}$ be a sequence of iid Bernoulli($\frac{1}{2}$) random variables. We let n_1 be the smallest integer such that the following holds

$$\Pr(2^{-2 \sum_{i=1}^{n_1} B_i} \leq \frac{P_e}{2^{n_0+n_1}}) \geq 1 - \frac{d}{3}. \quad (99)$$

It is easy to see that (99) is equivalent to

$$\Pr(\sum_{i=1}^{n_1} B_i \geq \log(\log \frac{1}{P_e}) + \log(n_0 + n_1)) \geq 1 - \frac{d}{3}. \quad (100)$$

Also, as the random variables B_i are Bernoulli($\frac{1}{2}$) and iid, the relation (100) is equivalent to

$$\frac{\sum_{j=0}^{\log(\log \frac{1}{P_e}) + \log(n_0+n_1)} \binom{n_1}{j}}{2^{n_1}} < \frac{d}{3}. \quad (101)$$

A sufficient condition for (101) to hold is as follows:

$$\frac{n_1^{1 + \log(\log \frac{1}{P_e}) + \log(n_0+n_1)}}{2^{n_1}} \leq \frac{d}{3},$$

and after applying the function $\log(\cdot)$ to both sides and some further simplifications we reach to

$$n_1 - (1 + \log(\log \frac{1}{P_e}) + \log(n_0 + n_1)) \log n_1 \geq \log \frac{3}{d}. \quad (102)$$

It can be shown through some simple steps that there are constants $c_6, c_7 > 0$ s.t. if we choose

$$n_1 = \left\lceil \log \frac{3}{d} + c_6 (\log(\log \frac{3}{d}))^2 + c_7 \log(\log(\frac{2}{P_e})) \log(\log \frac{3}{d}) \right\rceil, \quad (103)$$

then the inequality (102) holds. Now, let $\tilde{N} = 2^{n_0+n_1}$ and consider the set \mathcal{A}_1 defined as

$$\mathcal{A}_1 = \{i \in \{0, \dots, \tilde{N} - 1\} : Z(W_{\tilde{N}}^{(i)}) \leq \frac{P_e}{\tilde{N}}\}. \quad (104)$$

We now show that

$$\frac{|\mathcal{A}_1|}{\tilde{N}} \geq R. \quad (105)$$

This relation together with (104) shows that block error probability of the polar code of block-length \tilde{N} and rate R is at most P_e . In order to show (105), we consider the sub-channels \mathcal{A}_1 that are branched out from the ones in the set \mathcal{A} . Let $i \in \mathcal{A}$ and consider the sub-channel $W_{N_0}^{(i)}$. By using the relations (71), Lemma 20 and (99) we conclude the following. At level $n_0 + n_1$, the number of sub-channels that are branched out from $W_{N_0}^{(i)}$ and have Bhattacharyya value less than $\frac{P_e}{\tilde{N}}$ is at least

$$2^{n_1} (1 - c_2 Z(W_{N_0}^{(i)}) (1 + \log \frac{1}{Z(W_{N_0}^{(i)})})) (1 - \frac{d}{3}).$$

Hence, by using (98) the total number of sub-channels at level $n_0 + n_1$ that are branched out from a sub-channel in \mathcal{A} and have Bhattacharyya value less than $\frac{P_e}{N}$ is

$$2^{n_0+n_1} \left(R + \frac{2}{3}d\right) \left(1 - \frac{d}{3}\right) (1 - c_2 \sum_{i \in \mathcal{A}} Z(W_{N_0}^{(i)}) (1 + \log \frac{1}{Z(W_{N_0}^{(i)})})). \quad (106)$$

Now, by using Lemma 21 we have

$$\begin{aligned} & c_2 \sum_{i \in \mathcal{A}} Z(W_{N_0}^{(i)}) (1 + \log \frac{1}{Z(W_{N_0}^{(i)})}) \\ & \leq 2c_2 c_3 \sum_{i \in \mathcal{A}} (Z(W_{N_0}^{(i)}) (1 - Z(W_{N_0}^{(i)})))^\alpha \\ & \leq 2c_2 c_3 \mathbb{E}[(Z_{n_0} (1 - Z_{n_0}))^\alpha] \\ & \leq 2c_2 c_3 2^{-n_0 \rho} \\ & \stackrel{(96)}{\leq} \frac{d}{3}. \end{aligned}$$

Therefore, the expression (106) is lower-bounded by

$$2^{n_0+n_1} \left(R + \frac{2}{3}d\right) \left(1 - \frac{d}{3}\right)^2 \geq 2^{n_0+n_1} R = \tilde{N}R.$$

Hence, the relation (105) is proved and a block-length of size \tilde{N} is sufficient to achieve a rate R and error at most P_e . It is now easy to see that $\log \tilde{N} = n_0 + n_1$ has the form of (93).

ACKNOWLEDGMENT

The authors wish to thank Erdal Arıkan, Ali Goli, and Emre Telatar for their valuable comments on this topic.

REFERENCES

- [1] E. Arıkan, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," *IEEE Transactions on Information Theory*, vol. 55 (7), pp. 3051–3073, 2009.
- [2] R.G. Gallager, "A simple derivation of the coding theorem and some applications", *IEEE Transactions on Information Theory*, vol. 11, no. 1, pp. 3-18, 1965.
- [3] E. Arıkan and E. Telatar, "On the rate of channel polarization," in *Proc. ISIT*, Seoul, South Korea, pp.1493-1495, 2009.
- [4] S. H. Hassani, R. Mori, T. Tanaka and R. Urbanke, "Rate dependent analysis of the asymptotic behavior of channel polarization", *IEEE Transactions on Information Theory*, *IEEE Transactions on Information Theory*, vol. 59 (4) pp. 2267–2276, 2013.
- [5] R. L. Dobrushin, "Mathematical problems in the Shannon theory of optimal coding of information", in *Proc. 4th Berkeley Symp. Mathematics, Statistics, and Probability*, vol. 1, pp. 211-252, 1961.
- [6] V. Strassen, "Asymptotische abschätzungen in Shannon informations-theorie", in *Trans. 3d Prague Conf. Inf. Theory*, Prague, pp. 689-723, 1962.
- [7] Y. Polyanskiy, H. V. Poor, and S. Verdú, "A channel coding rate in the finite block-length regime", *IEEE Trans. Inf. Theory*, 56 (5), pp. 2307-2359, 2010.
- [8] S. B. Korada, "Polar codes for channel and source coding," Ph.D. dissertation, EPFL, Lausanne, Switzerland, July 2009.
- [9] R. Mori and T. Tanaka, "Performance and construction of polar codes on symmetric binary-input memoryless channels", in *Proc. ISIT*, Seoul, South Korea, pp.1496-1500, 2009.
- [10] S. H. Hassani, S. B. Korada and R. Urbanke, "The compound capacity of polar codes", in *proc. 47th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pp.16-21, 2009.
- [11] S. B. Korada, A. Montanari, E. Telatar and R. Urbanke, "An empirical scaling law for polar codes", in *Proc. ISIT*, Texas, USA, pp.884-888, 2010.
- [12] S. H. Hassani, K. Alishahi and R. Urbanke, "On the scaling of polar codes: II. The behavior of un-polarized channels", in *Proc. ISIT*, Texas, USA, pp.879-883, 2010.

- [13] T. Richardson and R. Urbanke, *Modern Coding Theory*. Cambridge University Press, 2008.
- [14] Chris Godsil and Gordon Royle, *Algebraic Graph Theory*. Springer, 2001.
- [15] I. Tal and A. Vardy, "How to construct polar codes," presented at 2010 IEEE Info. Theory Workshop, Dublin, Ireland, 2010. [online] Available: arXiv:1105.6164v1 [cs.IT].
- [16] R. Pedarsani, H. Hassani, I. Tal and E. Telatar, "On the construction of polar codes," in *Proc. ISIT*, St. Petersburg, Russia, pp. 11–15, 2011.
- [17] See http://en.wikipedia.org/wiki/Sturms_theorem
- [18] M. Bastani Parizi and E. Telatar, "On Correlation Between Polarized BECs," [online] Available: arXiv:1301.5536 [cs.IT].
- [19] V. Guruswami, P. Xia, "Polar codes: Speed of polarization and polynomial gap to capacity," in *Electronic Colloquium on Computational Complexity*, Report No. 50, 2013.

APPENDIX A PROOFS

1) *Proof of Lemma 5:* The proof of the right side (59) and also (60) is an easy application of the Markov inequality. To prove the left side of (59), we define sequences $\{x_n\}_{n \geq 1}$ and $\{y_n\}_{n \geq 1}$ as

$$x_n = 2^{-n}, \quad (107)$$

$$y_n = 1 - 2^{-n}. \quad (108)$$

We start by noting that

$$\begin{aligned} \mathbb{E}(Z_n(1 - Z_n)) & \leq \sum_{i=1}^n 2^{-i} \Pr(Z_n \in [x_{i+1}, x_i]) \\ & \quad + \sum_{i=1}^n 2^{-i} \Pr(Z_n \in [y_i, y_{i+1}]) \\ & \quad + 2^{-n}. \end{aligned}$$

As a result, there exists an index $j \in \{1, \dots, n\}$ such that at least one of the following cases occurs:

$$\mathbb{E}[Z_n(1 - Z_n)] \leq 2n[2^{-j} \Pr(Z_n \in [x_{j+1}, x_j]) + 2^{-n}], \quad (109)$$

or

$$\mathbb{E}[Z_n(1 - Z_n)] \leq 2n[2^{-j} \Pr(Z_n \in [y_j, y_{j+1}]) + 2^{-n}]. \quad (110)$$

We show that in each of these cases the statement of the lemma holds. Firstly, note that because of the symmetry of Z_n we can write

$$\Pr(Z_n^z \in [y_{j+1}, y_j]) = \Pr(Z_n^{1-z} \in [x_{j+1}, x_j]).$$

Hence, without loss of generality we can assume that (109) holds. We first prove the lemma for $a = 1 - b = \frac{1}{4}$. We then use this result to prove the lemma in its fullest extent. We claim that for any $1 \leq j \leq n$ we have,

$$2^{-j} \Pr(Z_n \in [x_{j+1}, x_j]) \leq 2(n+1) \Pr(Z_n \in [\frac{1}{4}, \frac{3}{4}]) + \frac{n^3}{2^n}. \quad (111)$$

Assuming that the above claim holds true, by using (109) we obtain

$$\mathbb{E}(Z_n(1 - Z_n)) \leq 2n \left[\Pr(Z_n \in [\frac{1}{4}, \frac{3}{4}]) + \frac{n^2 + 2}{2^n} \right],$$

and as a result, by taking $\frac{1}{n} \log(\cdot)$ from both sides, the first part of the lemma is proved for $a = 1 - b = \frac{1}{4}$.

We now turn to the proof of relation (111) for $1 \leq j \leq n$. For $j = 1$, the result of the claim is trivial. Hence, in the following

we assume that $2 \leq j \leq n$. We now prove that for any fixed j such that $2 \leq j \leq n$, we have

$$2^{-j} \Pr(Z_n \in [x_{j+1}, x_j]) \leq 2(n+1) \Pr(Z_n \in [\frac{1}{4}, \frac{3}{4}]) + \frac{n^3}{2^n}, \quad (112)$$

and hence the relation (111) is also proved. We fix the index j and prove the above claim for any value of $n \in \mathbb{N}$. The proof consist of two steps.

Step1: We first show that $\forall m \in \mathbb{N}$,

$$\Pr(Z_m \in [x_{2j+2}, x_j]) \leq m \Pr(Z_m \in [x_j, \frac{3}{4}]) + \frac{1}{2^n}. \quad (113)$$

To prove (113), fix $m \in \mathbb{N}$ and define the sets A and B as

$$A = \{(b_1, \dots, b_m) \in \Omega_m : t_{b_m} \circ \dots \circ t_{b_1}(z) \in [x_{2j+2}, x_j]\}.$$

$$B = \{(b_1, \dots, b_m) \in \Omega_m : t_{b_m} \circ \dots \circ t_{b_1}(z) \in [x_j, \frac{3}{4}]\}.$$

In other words, A is the set of all the paths that start from $z = Z_0$ and end up in $[x_{2j+2}, x_j]$ and B is the set of paths that start from z and end up in $[x_j, \frac{3}{4}]$. We now partition the A into the disjoint sets A_k , $k \in \{0, 1, \dots, m\}$, defined as

$$A_k = \{(b_1, \dots, b_m) \in A : b_k = 1; b_i = 0 \ \forall i > k\}. \quad (114)$$

It is easy to see that $|A - \cup_k A_k| \leq 1$. Our aim is now to show that for $k \in \{0, 1, \dots, m\}$,

$$|A_k| \leq |B|. \quad (115)$$

To do this, we show that there exists a one-to-one correspondence between A_k and a subset of B . In other words, we claim that we can map each member of A_k to a distinct member of B . Consider $(b_1, \dots, b_m) \in A_k$. We now construct a distinct member $(b'_1, \dots, b'_m) \in B$ corresponding to (b_1, \dots, b_m) . We first set $b'_i = b_i$ for $i < k$ and hence the uniqueness condition is fulfilled. Consider the number x defined as

$$x = \begin{cases} z & \text{if } k = 1, \\ t_{b_{k-1}} \circ \dots \circ t_{b_1}(z) & \text{if } k > 1. \end{cases} \quad (116)$$

Note that since $(b_1, \dots, b_m) \in A_k$ we have

$$t_{b_m} \circ \dots \circ t_{b_k}(x) \in [x_{2j+1}, x_j]. \quad (117)$$

Now, note that as $(b_1, \dots, b_m) \in A_k$, we have $b_k = 1$ and $b_i = 0$ for $i > k$. Thus, in this setting (117) becomes

$$\overbrace{t_0 \circ \dots \circ t_0}^{m-k \text{ times}}(x^2) \in [x_{2j+1}, x_j]. \quad (118)$$

Hence,

$$x_{2j+1} \leq 1 - (1 - x^2)^{2^{m-k}} \leq x_j. \quad (119)$$

From the left side of (119) and using the fact that $1 - (1 - x)^2 \leq 2x$ we obtain

$$x_{2j+1} \leq 2^{m-k} x^2 \Rightarrow 2^{-j+k-m} \leq x. \quad (120)$$

From the right side of (119) we have

$$\ln(1 - x_j) \leq 2^{m-k} \ln(1 - x^2),$$

and by using the inequality $-x - \frac{x^2}{2} \leq \ln(1 - x) \leq -x$ we obtain

$$x \leq 2^{\frac{-j+k-m+1}{2}}. \quad (121)$$

Let us recall that we let $b'_i = b_i$ for $i < k$. We now construct the remaining values b'_k, \dots, b'_m by the following algorithm: consider the number x given in (116). In the following, we will also construct a sequence $x = x_{k-1}, x_k, x_{k+1}, \dots, x_m$ such that for $i \geq k$ we have $x_i = t_{b'_i}(x_{i-1})$. Begin with the initial value $x_{k-1} = x$ and for $i \geq k$ recursively construct b'_i from b'_{i-1} and x_{i-1} by the following rule: if $t_{b'_{i-1}}(x_{i-1}) \leq \frac{3}{4}$, then $b'_i = 0$ and $x_i = t_0(x_{i-1})$, otherwise $b'_i = 1$ and $x_i = t_1(x_{i-1})$. We now show that the value of x_m is always in the interval $[x_j, \frac{3}{4}]$. In this regard, an important observation is that for i s.t. $k-1 \leq i \leq m$, once the value of x_i lies in the interval $[x_j, \frac{3}{4}]$ then for all $i \leq t \leq m$ we have $x_t \in [x_j, \frac{3}{4}]$. Hence, we only need to show that by the above algorithm, there exists an index i , s.t. $k-1 \leq i \leq m$, and the value of x_i lies in the interval $[x_j, \frac{3}{4}]$. On one hand, observe that due to (121) and the fact that $j \geq 2$, we have $x \leq 2^{-\frac{1}{2}} < \frac{3}{4}$. Thus, the value of x_i is definitely less than $\frac{3}{4}$ for $i \geq k$. If the value of x_{k-1} is also greater than x_j then we have nothing to prove. Else, it might be the case that $x < x_j$. We now show that in this case the algorithm moves in a way that the value of x_m falls eventually into the desired region $[x_j, \frac{3}{4}]$. To show this, a moment of thought reveals that this is equivalent to showing that we always have

$$\overbrace{t_0 \circ \dots \circ t_0}^{m-k+1 \text{ times}}(x) = 1 - (1 - x)^{2^{m-k+1}} \geq x_j. \quad (122)$$

Note that the function $1 - (1 - x)^{2^{m-k+1}}$ is a strictly increasing function of the unit interval. Thus, in order to have (122) it is equivalent that

$$2^{m-k+1} \ln(1 - x) \leq \ln(1 - x_j),$$

and after some further simplification using the inequality $-x - \frac{x^2}{2} \leq \ln(1 - x) \leq -x$, we deduce that a sufficient condition to have (122) is

$$x_j \leq 2^{m-k} x \Rightarrow 2^{-j+k-m} \leq x. \quad (123)$$

But this sufficient condition is certainly met by considering the inequality (120) and noting the fact that $-j + \frac{k-m+1}{2} \geq -j + k - m$. Hence, the claim in (115) is proved and as a result, the claim in (113) is true.

Step 2: Firstly note that in order for Z_n to be in the interval $[x_{j+1}, x_j]$, the value of Z_{n-j} should lie in the interval $[x_{2j+1}, x_j^{2^{-2^j}}]$. As a result, we can write

$$\begin{aligned} & \Pr(Z_n \in [x_{j+1}, x_j]) \\ &= \Pr(Z_n \in [x_{j+1}, x_j] | Z_{n-j} \in [x_{2j+1}, x_j]) \\ & \quad \times \Pr(Z_{n-j} \in [x_{2j+1}, x_j]) \\ & \quad + \Pr(Z_n \in [x_{j+1}, x_j] | Z_{n-j} \in (x_j, x_j^{2^{-2^j}}]) \\ & \quad \times \Pr(Z_{n-j} \in (x_j, x_j^{2^{-2^j}}]), \end{aligned} \quad (124)$$

and by letting $m = n - j$ in relation (113), we can easily obtain

$$\Pr(Z_{n-j} \in [x_{2j+1}, x_j]) \leq n \Pr(Z_{n-j} \in [x_j, \frac{3}{4}]) + \frac{n^2 + 1}{2^n}. \quad (125)$$

Thus, by combining (124) and (125), we obtain

$$\Pr(Z_n \in [x_{j+1}, x_j])$$

$$\leq n\Pr(Z_{n-j} \in [x_j, \frac{3}{4}]) + \Pr(Z_{n-j} \in [x_j, x_j^{2^{-2j}}]) + \frac{n^2 + 1}{2^n}. \quad (126)$$

Finally, in order to conclude the proof of (112), we prove the following relations:

$$2^{-j}\Pr(Z_{n-j} \in [x_j, \frac{3}{4}]) \leq \Pr(Z_n \in [\frac{1}{4}, \frac{3}{4}]), \quad (127)$$

and

$$2^{-j}\Pr(Z_{n-j} \in [x_j, x_j^{2^{-2j}}]) \leq \Pr(Z_n \in [\frac{1}{4}, \frac{3}{4}]). \quad (128)$$

Firstly note that since $(x_j)^{\frac{1}{2^{2j}}} \geq \frac{3}{4}$, then it is enough to prove (128). To prove (128), we only need to show that for a value x s.t. $x \in [x_j, (x_j)^{\frac{1}{2^{2j}}}]$, there exists an j -tuple $(b_1, \dots, b_j) \in \Omega_j$ such that $t_{b_1} \circ \dots \circ t_{b_j}(x) \in [\frac{1}{4}, \frac{3}{4}]$. We show this by constructing the binary values b_1, \dots, b_j in terms of x . Consider the following algorithm: start with $y_0 = x$ and for $1 \leq i \leq j$, we recursively construct b_i from y_{i-1} by the following rule: If $t_0(y_{i-1}) \leq \frac{3}{4}$, then $b_i = 0$ and $y_i = t_0(y_{i-1})$. Otherwise, let $b_i = 1$ and $y_i = t_1(x_{i-1})$. To show that this algorithm succeeds in the sense that $y_j \in [\frac{1}{4}, \frac{3}{4}]$, we first observe that once the value of y_i lies in the interval $[\frac{1}{4}, \frac{3}{4}]$ (for some $1 \leq i \leq j$), then for all $i \leq t \leq j$ we have $y_t \in [\frac{1}{4}, \frac{3}{4}]$. Hence, we only need to show that by the above algorithm, there exists an index i , s.t. $1 \leq i \leq j$, and the value of y_i lies in the interval $[\frac{1}{4}, \frac{3}{4}]$. On one hand, assume $y_0 = x \in [x_j, \frac{1}{4}]$. We can then write

$$\begin{aligned} \overbrace{t_0 \circ \dots \circ t_0}^{j \text{ times}}(x) &= 1 - (1 - x)^{2^j} \\ &\geq 1 - (1 - x_j)^{2^j} \\ &\geq \frac{1}{2}, \end{aligned}$$

where the last step follows from the fact that $x_j = 2^{-j}$. On the other hand, assume $x \in (\frac{3}{4}, (x_j)^{\frac{1}{2^j}}]$. We can write

$$\begin{aligned} \overbrace{t_1 \circ \dots \circ t_1}^{j \text{ times}}(x) &\leq ((x_j)^{\frac{1}{2^{2j}}})^{2^{2j}} \\ &\leq x_j < \frac{3}{4}. \end{aligned}$$

As a result, the above algorithm always succeeds and the lemma is proved for $a = 1 - b = \frac{1}{4}$.

We now prove the lemma for any choice of $a, b \in (0, 1)$ s.t. $\sqrt{a} \leq 1 - \sqrt{1 - b}$. Let $p_n(z, a, b)$ be defined as in (27). We have

$$\begin{aligned} p_{n+1}(z, a, b) &= \sum_{\phi_{\omega_{n+1}}} \frac{1}{2^{n+1}} \mathbb{1}_{\{z \in \phi_{\omega_{n+1}}^{-1}[a, b]\}} \\ &= \sum_{\phi_{\omega_n}} \frac{1}{2^n} \frac{\mathbb{1}_{\{z \in \phi_{\omega_n}^{-1}[t_0^{-1}(a), t_0^{-1}(b)]\}} + \mathbb{1}_{\{z \in \phi_{\omega_n}^{-1}[t_1^{-1}(a), t_1^{-1}(b)]\}}}{2} \\ &= \frac{1}{2} (p_n(z, t_0^{-1}(a), t_0^{-1}(b)) + p_n(z, t_1^{-1}(a), t_1^{-1}(b), z)). \end{aligned}$$

It is easy to see that if $\sqrt{a} \leq 1 - \sqrt{1 - b}$, then

$$[t_0^{-1}(a), t_1^{-1}(b)] \subseteq [t_0^{-1}(a), t_0^{-1}(b)] \cup [t_1^{-1}(a), t_1^{-1}(b)],$$

and hence,

$$2p_{n+1}(z, a, b) \geq p_n(z, t_0^{-1}(a), t_1^{-1}(b)).$$

Continuing this way, we can show that for $m \in \mathbb{N}$

$$\begin{aligned} 2^m p_{n+m}(z, a, b) &\geq p_n(z, \overbrace{t_0^{-1} \circ \dots \circ t_0^{-1}}^{m \text{ times}}(a), \overbrace{t_1^{-1} \circ \dots \circ t_1^{-1}}^{m \text{ times}}(b)). \end{aligned} \quad (129)$$

As m grows large, we have

$$\begin{aligned} \overbrace{t_0^{-1} \circ \dots \circ t_0^{-1}}^{m \text{ times}}(a) &\rightarrow 0, \\ \overbrace{t_1^{-1} \circ \dots \circ t_1^{-1}}^{m \text{ times}}(b) &\rightarrow 1. \end{aligned}$$

Therefore, by (129) there exists a positive integer m_0 such that for $n \in \mathbb{N}$

$$2^{m_0} p_{n+m_0}(z, a, b) \geq p_n(z, \frac{1}{4}, \frac{3}{4}).$$

The thesis now follows from this relation together with the result of Lemma 54.

2) *Proof of Lemma 10:* Recall that for a realization $\omega = \{b_k\}_{k \in \mathbb{N}} \in \Omega$ we define $\omega_n = (b_1, \dots, b_n)$. The maps t_0 and t_1 , hence the maps ϕ_{ω_n} s, are strictly increasing maps on $[0, 1]$. Thus $\phi_{\omega_n}(z) \rightarrow 0$ implies that $\phi_{\omega_n}(z') \rightarrow 0$ for $z' \leq z$ and $\phi_{\omega_n}(z) \rightarrow 1$ implies that $\phi_{\omega_n}(z') \rightarrow 1$ for $z' \geq z$. Moreover, we know that for almost every $z \in (0, 1)$, $\lim_{n \rightarrow \infty} \phi_{\omega_n}(z)$ is either 0 or 1 for almost every realization $\{\phi_{\omega_n}\}_{n \in \mathbb{N}}$. Hence, it suffices to let

$$z_{\omega}^* = \inf\{z : \phi_{\omega_n}(z) \rightarrow 1\}.$$

To prove the second part of the lemma, notice that

$$\begin{aligned} z &= \Pr(Z_{\infty} = 1) \\ &= \Pr(\phi_{\omega_n}(z) \rightarrow 1) \\ &= \Pr(\inf\{z : \phi_{\omega_n}(z) \rightarrow 1\} \leq z) \\ &= \Pr(z_{\omega}^* < z). \end{aligned}$$

Which shows that z_{ω}^* is uniformly distributed on $[0, 1]$.

3) *Proof of Lemma 11:* In order to compute $\lim_{n \rightarrow \infty} \mathbb{E}[\frac{1}{n} \log(\phi_{\omega_n}^{-1}(b) - \phi_{\omega_n}^{-1}(a))]$, we first define the process $\{\bar{Z}_n\}_{n \in \mathbb{N} \cup \{0\}}$ with $\bar{Z}_0 = z \in [0, 1]$ and

$$\bar{Z}_{n+1} = \begin{cases} \sqrt{\bar{Z}_n}, & \text{w.p. } \frac{1}{2}, \\ 1 - \sqrt{1 - \bar{Z}_n}, & \text{w.p. } \frac{1}{2}. \end{cases} \quad (130)$$

We can think of \bar{Z}_n as the reverse stochastic process of Z_n . Equivalently, we can also define \bar{Z}_n via the inverse maps t_0^{-1}, t_1^{-1} . Consider the sequence of i.i.d. symmetric Bernoulli random variables B_1, B_2, \dots and define $\bar{Z}_n = \psi_{\omega_n}(z)$ where $\omega_n \triangleq (b_1, \dots, b_n) \in \Omega_n$ and

$$\psi_{\omega_n} = t_{b_n}^{-1} \circ t_{b_{n-1}}^{-1} \circ \dots \circ t_{b_1}^{-1}. \quad (131)$$

We now show that the Lebesgue measure (or the uniform probability measure) on $[0, 1]$, denoted by ν , is the unique, hence ergodic, invariant measure for the Markov process

\bar{Z}_n . To prove this result, first note that if \bar{Z}_n is distributed according to the Lebesgue measure, then

$$\begin{aligned}\Pr(\bar{Z}_{n+1} < x) &= \frac{1}{2}\Pr(\bar{Z}_n < t_0(x)) + \frac{1}{2}\Pr(\bar{Z}_n < t_1(x)) \\ &= \frac{1}{2}x^2 + \frac{1}{2}(2x - x^2) = x.\end{aligned}$$

Thus, \bar{Z}_{n+1} is also distributed according to the Lebesgue measure and this implies the invariance of the Lebesgue measure for \bar{Z}_n . In order to prove the uniqueness, we will show that for any $z \in (0, 1)$, \bar{Z}_n converges weakly to a uniformly distributed random point in $[0, 1]$, i.e.,

$$\bar{Z}_n = \psi_{\omega_n}(z) \xrightarrow{d} \nu. \quad (132)$$

Note that with (132) the uniqueness of ν is proved since for any invariant measure ρ assuming \bar{Z}_n is distributed according to ρ , we have

$$\rho(\cdot) = \Pr(\bar{Z}_n \in \cdot) = \int \Pr(\bar{Z}_n \in \cdot) \rho(dz) \xrightarrow{d} \nu(\cdot). \quad (133)$$

To prove (132), note that ψ_{ω_n} has the same (probability) law as $\phi_{\omega_n}^{-1}$ and we know that $\phi_{\omega_n}^{-1}(z) \rightarrow z_\omega^*$ almost surely and hence weakly. Also, z_ω^* is distributed according to ν , which proves (132). We are now ready to show that

$$\lim_{n \rightarrow \infty} \mathbb{E}\left[\frac{1}{n} \log(\phi_{\omega_n}^{-1}(b) - \phi_{\omega_n}^{-1}(a))\right] = \frac{1}{2 \ln 2} - 1. \quad (134)$$

Using the mean value theorem, we can write

$$\psi_n(a) - \psi_n(b) = \psi'_n(c)(b - a),$$

for some $c \in (a, b)$. And by chain rule,

$$\begin{aligned}\psi'_{\omega_n}(c) &= (t_{b_n}^{-1} \circ t_{b_{n-1}}^{-1} \circ \dots \circ t_{b_1}^{-1})'(c) \\ &= t_{b_1}^{-1'}(c) \cdot t_{b_2}^{-1'}(t_{b_1}^{-1}(c)) \cdot \dots \cdot t_{b_n}^{-1'}(t_{b_{n-1}}^{-1} \circ \dots \circ t_{b_1}^{-1}(c)) \\ &= t_{b_1}^{-1'}(\psi_0(c)) \cdot t_{b_2}^{-1'}(\psi_1(c)) \cdot \dots \cdot t_{b_n}^{-1'}(\psi_{n-1}(c)),\end{aligned}$$

and after applying $\log(\cdot)$ to both sides we obtain

$$\frac{1}{n} \log(\psi'_{\omega_n}(c)) = \frac{1}{n} \sum_{j=1}^n \ln t_{b_j}^{-1'}(\psi_{j-1}(c)). \quad (135)$$

By the ergodic theorem, the last expression converges almost surely to the expectation of $\log t_{B_1}^{-1'}(U)$, where U is assumed to be distributed according to ν . Hence, the asymptotic value of (135) can be computed as

$$\begin{aligned}\mathbb{E}[\log t_{B_1}^{-1'}(U)] &= \frac{1}{2} \int_0^1 \log(\sqrt{x})' dx + \frac{1}{2} \int_0^1 \log(1 - \sqrt{1-x})' dx \\ &= \frac{1}{2 \ln 2} - 1.\end{aligned}$$

APPENDIX B AUXILIARY LEMMAS

Lemma 19: Consider a channel W with its Bhattacharyya process $Z_n = Z(W_n)$ and assume that for $n \in \mathbb{N}$

$$\mathbb{E}[(Z_n(1 - Z_n))^\alpha] \leq \beta 2^{-n\rho}, \quad (136)$$

where α, β, ρ are positive constants with $\alpha < 1$. We then have for $n \in \mathbb{N}$

$$\Pr(Z_n \leq \frac{1}{2}) \geq I(W) - c_1 2^{-n\rho}, \quad (137)$$

where c_1 is a positive constant that depends on α, β, ρ .

Proof: The proof consists of three steps. First, consider an arbitrary BMS channel W and let $Z_n = Z(W_n)$. Also, consider the process $E_n = 1 - Z_n^2$. By using the relations (12) and (13), it can easily be checked that the process E_n has the form of (140) and hence Lemma 20 is applicable to E_n . We thus have from (141) that for $n \in \mathbb{N}$

$$\Pr(E_n \geq \frac{1}{2}) \leq c_2 E_0 (1 + \log \frac{1}{E_0}).$$

As a consequence

$$\begin{aligned}I(W) &= \lim_{n \rightarrow \infty} \Pr(E_n \geq \frac{1}{2}) \\ &\leq c_2 (1 - Z(W)^2) (1 + \log \frac{1}{1 - Z(W)^2}).\end{aligned} \quad (138)$$

In the second step, we consider a channel W for which (136) holds for $n \in \mathbb{N}$. By using (136), it is easy to see that for $n \in \mathbb{N}$

$$\begin{aligned}\mathbb{E}[(Z_n^2(1 - Z_n^2))^\alpha \mathbb{1}_{\{Z_n \geq \frac{1}{2}\}}] &= \mathbb{E}[(Z_n(1 + Z_n))^\alpha (Z_n(1 - Z_n))^\alpha \mathbb{1}_{\{Z_n \geq \frac{1}{2}\}}] \\ &\leq \sup_{z \in [\frac{1}{2}, 1]} (z(1+z))^\alpha \mathbb{E}[(Z_n(1 - Z_n))^\alpha \mathbb{1}_{\{Z_n \geq \frac{1}{2}\}}] \\ &\leq 2^\alpha \beta 2^{-n\rho} \leq \beta 2^{1-n\rho}.\end{aligned} \quad (139)$$

In the final step, we consider a number $n \in \mathbb{N}$ and let $N = 2^n$. We then define the set \mathcal{A} as

$$\mathcal{A} = \{i \in \{0, 1, \dots, N-1\} : Z(W_N^{(i)}) \leq \frac{1}{2}\},$$

with \mathcal{A}^c being its complement. We have

$$\begin{aligned}\sum_{i \in \mathcal{A}^c} I(W_N^{(i)}) &\stackrel{(a)}{\leq} \sum_{i \in \mathcal{A}^c} c_2 (1 - Z(W_N^{(i)})^2) (1 + \log \frac{1}{1 - Z(W_N^{(i)})^2}) \\ &\stackrel{(b)}{\leq} \sum_{i \in \mathcal{A}^c} 4c_2 c_3 (Z(W_N^{(i)})^2 (1 - Z(W_N^{(i)})^2))^\alpha \\ &= 4c_2 c_3 N \mathbb{E}[(Z_n^2(1 - Z_n^2))^\alpha \mathbb{1}_{\{Z_n \geq \frac{1}{2}\}}] \\ &\stackrel{(c)}{\leq} 8c_2 c_3 N \beta 2^{-n\rho}.\end{aligned}$$

Here (a) follows from (138), (b) follows from Lemma 21 and the fact that for $x \leq \frac{3}{4}$ we have $1 + \log \frac{1}{x} \leq 4 \log \frac{1}{x}$, and (c) follows from (139). Now, as a consequence of the above chain of inequalities we have

$$\begin{aligned}|\mathcal{A}| &\geq \sum_{i \in \mathcal{A}} I(W_N^{(i)}) \\ &= NI(W) - \sum_{i \in \mathcal{A}^c} I(W_N^{(i)}) \\ &\geq N(I(W) - 2c_2 c_3 \beta 2^{-n\rho}),\end{aligned}$$

and consequently

$$\Pr(Z_n \leq \frac{1}{2}) = \frac{|\mathcal{A}|}{N} \geq 2c_2 c_3 \beta 2^{-n\rho}.$$

Hence, the proof follows. ■

Lemma 20: Consider a generic stochastic process $\{X_n\}_{n \geq 0}$ s.t. $X_0 = x$, where $x \in (0, 1)$ and for $n \geq 1$

$$X_n \leq \begin{cases} X_{n-1}^2 & \text{if } B_n = 1, \\ 2X_{n-1} & \text{if } B_n = 0. \end{cases} \quad (140)$$

Here, $\{B_n\}_{n \geq 1}$ is a sequence of iid random variables with distribution Bernoulli($\frac{1}{2}$). We then have for $n \in \mathbb{N}$

$$\Pr(X_n \leq 2^{-2^{\sum_{i=1}^n B_i}}) \geq 1 - c_2 x (1 + \log \frac{1}{x}), \quad (141)$$

where c_2 is a positive constant.

Proof: We analyze the process $A_n = -\log X_n$, i.e., $A_0 = -\log x \triangleq a_0$ and

$$A_{n+1} = \begin{cases} 2A_n & \text{if } B_n = 1, \\ A_n - 1 & \text{if } B_n = 0. \end{cases} \quad (142)$$

Note that in terms of the process A_n , the statement of the lemma can be phrased as

$$\Pr(A_n \geq 2^{\sum_{i=1}^n B_i}) \geq 1 - c_2 \frac{1 + a_0}{2^{a_0}}.$$

Associate to each $(b_1, \dots, b_n) \triangleq \omega_n \in \Omega_n$ a sequence of “runs” $(r_1, \dots, r_{k(\omega_n)})$. This sequence is constructed by the following procedure. We define r_1 as the smallest index $i \in \mathbb{N}$ so that $b_{i+1} \neq b_1$. In general, if $\sum_{j=1}^{k-1} r_j < n$ then

$$r_k = \min\{i \mid \sum_{j=1}^{k-1} r_j < i \leq n, b_{i+1} \neq b_{\sum_{j=1}^{k-1} r_j}\} - \sum_{j=1}^{k-1} r_j.$$

The process stops whenever the sum of the runs equals n . Denote the stopping time of the process by $k(\omega_n)$. In words, the sequence (b_1, \dots, b_n) starts with b_1 . It then repeats b_1 , r_1 times. Next follow r_2 instances of \bar{b}_1 ($\bar{b}_1 := 1 - b_1$), followed again by r_3 instances of b_1 , and so on. We see that b_1 and $(r_1, \dots, r_{k(\omega_n)})$ fully describe $\omega_n = (b_1, \dots, b_n)$. Therefore, there is a one-to-one map

$$(b_1, \dots, b_n) \longleftrightarrow \{b_1, (r_1, \dots, r_{k(\omega_n)})\}. \quad (143)$$

Note that we can either have $b_1 = 1$ or $b_1 = 0$. We start with the first case, i.e., we first assume $B_1 = 1$. We have:

$$\sum_{i=1}^n b_i = \sum_{j \text{ odd} \leq k(\omega_n)} r_j,$$

and

$$n = \sum_{j=1}^{k(\omega_n)} r_j.$$

Analogously, for a realization $(b_1, b_2, \dots) \triangleq \omega \in \Omega$ of the infinite sequence of random variable $\{B_i\}_{i \in \mathbb{N}}$, we can associate a sequence of runs (r_1, r_2, \dots) . In this regard, considering the infinite sequence of random variables $\{B_i\}_{i \in \mathbb{N}}$ (with the extra condition $B_1 = 1$), the corresponding sequence of runs, which we denote by $\{R_k\}_{k \in \mathbb{N}}$, is an iid sequence with $\Pr(R_i = j) = \frac{1}{2^j}$. Let us now see how we can express the A_n in terms of the $r_1, r_2, \dots, r_{k(\omega_n)}$. We begin by a simple example: Consider the sequence $(b_1 = 1, b_2, \dots, b_8)$ and the associated run sequence $(r_1, \dots, r_5) = (1, 2, 1, 3, 1)$. We have

$$\begin{aligned} A_1 &= a_0 2^{r_1}, \\ A_3 &= a_0 2^{r_1} - r_2, \end{aligned}$$

$$\begin{aligned} A_4 &= (a_0 2^{r_1} - r_2) 2^{r_3} = a_0 2^{r_1+r_3} - r_2 2^{r_3}, \\ A_7 &= (a_0 2^{r_1} - r_2) 2^{r_3} - r_4 = a_0 2^{r_1+r_3} - r_2 2^{r_3} - r_4, \\ A_8 &= ((a_0 \times 2^{r_1} - r_2) \times 2^{r_3} - r_4) \times 2^{r_5} \\ &= a_0 2^{r_1+r_3+r_5} - r_2 2^{r_3+r_5} - r_4 2^{r_5} \\ &= 2^{r_1+r_3+r_5} (a_0 - 2^{-r_1} r_2 - 2^{-(r_1+r_3)} r_4). \end{aligned}$$

In general, for a sequence (b_1, \dots, b_n) with the associated run sequence $(r_1, \dots, r_{k(\omega_n)})$ we can write:

$$\begin{aligned} A_n &= a_0 2^{\sum_{i \text{ odd} \leq k(\omega_n)} r_i} - \sum_{i \text{ even} \leq k(\omega_n)} r_i 2^{\sum_{j \text{ odd} < i} r_j} \\ &= a_0 2^{\sum_{i \text{ odd} \leq k(\omega_n)} r_i} - \sum_{i \text{ even} \leq k(\omega_n)} r_i 2^{(-\sum_{j \text{ odd} < i} r_j + \sum_{i \text{ odd} \leq k(\omega_n)} r_i)} \\ &= [2^{\sum_{i \text{ odd} \leq k(\omega_n)} r_i}] [a_0 - (\sum_{i \text{ even} \leq k(\omega_n)} r_i 2^{-\sum_{j \text{ odd} < i} r_j})] \\ &= [2^{\sum_{i=1}^n B_i}] [a_0 - (\sum_{i \text{ even} \leq k(\omega_n)} r_i 2^{-\sum_{j \text{ odd} < i} r_j})]. \end{aligned}$$

Our aim is to lower-bound

$$\begin{aligned} \Pr(A_n \geq 2^{\sum_{i=1}^n B_i}) \\ = \Pr(a_0 - \sum_{i \text{ even} \leq k(\omega_n)} r_i 2^{-\sum_{j \text{ odd} < i} r_j} \geq 1), \end{aligned}$$

or, equivalently, to upper-bound

$$\Pr(\sum_{i \text{ even} \leq k(\omega_n)} r_i 2^{-\sum_{j \text{ odd} < i} r_j} \geq a_0 - 1). \quad (144)$$

For $n \in \mathbb{N}$, define the set $U_n \in \mathcal{F}_n$ as

$$U_n = \{\omega_n \in \Omega_n \mid \exists l \leq k(\omega_n) : \sum_{i \text{ even} \leq l} r_i 2^{-\sum_{j \text{ odd} < i} r_j} \geq a_0 - 1\}.$$

Clearly we have:

$$\Pr(\sum_{i \text{ even} \leq k(\omega_n)} r_i 2^{-\sum_{j \text{ odd} < i} r_j} \geq a_0 - 1) \leq \Pr(U_n).$$

In the following we show that if $(b_1, \dots, b_n) \in U_n$, then for any choice of b_{n+1} , $(b_1, \dots, b_n, b_{n+1}) \in U_{n+1}$. We will only consider the case when $b_n, b_{n+1} = 1$, the other three cases can be verified similarly. Let $\omega_n = (b_1, \dots, b_{n-1}, b_n = 1) \in U_n$. Hence, $k(\omega_n)$ is an odd number (recall that $b_1 = 1$) and the quantity $\sum_{i \text{ even} \leq k(\omega_n)} r_i 2^{-\sum_{j \text{ odd} < i} r_j}$ does not depend on $r_{k(\omega_n)}$. Now consider the sequence $\omega_{n+1} = (b_1, \dots, b_n = 1, 1)$. Since the last bit (b_{n+1}) equals 1, then $r_{k(\omega_{n+1})} = r_{k(\omega_n)}$ and the value of the sum remains unchanged. As a result $(b_1, \dots, b_n, 1) \in U_{n+1}$. From above, we conclude that $\theta_i(U_i) \subseteq \theta_{i+1}(U_{i+1})$ and as a result

$$\Pr(U_i) = \Pr(\theta_i(U_i)) \leq \Pr(\theta_{i+1}(U_{i+1})) = \Pr(U_{i+1}).$$

Hence, the quantity $\lim_{n \rightarrow \infty} \Pr(U_n) = \lim_{n \rightarrow \infty} \Pr(\theta_n(U_n)) = \lim_{n \rightarrow \infty} \Pr(\cup_{i=1}^n \theta_i(U_i))$ is an upper bound on (144). On the other hand, consider the set

$$V = \{\omega \in \Omega \mid \exists l : \sum_{i \text{ even} \leq l} r_i 2^{-\sum_{j \text{ odd} < i} r_j} \geq a_0 - 1\}.$$

By the definition of V we have $\cup_{i=1}^{\infty} \theta_i(U_i) \subseteq V$, and as a result, $\Pr(\cup_{i=1}^{\infty} \theta_i(U_i)) \leq \Pr(V)$. In order to bound the probability of the set V , note that assuming $B_1 = 1$, the sequence $\{R_k\}_{k \in \mathbb{N}}$ (i.e., the sequence of runs when associated with the sequence

$\{B_i\}_{i \in \mathbb{N}}$ is an iid sequence with $\Pr(R_i = j) = \frac{1}{2^j}$. We also have

$$\begin{aligned} & \Pr(a_0 - \sum_{i \text{ even } \leq m} R_i 2^{-\sum_{j \text{ odd } < i} R_j} \leq 1) \\ &= \Pr(\sum_{i \text{ even } \leq m} R_i 2^{-\sum_{j \text{ odd } < i} R_j} \geq a_0 - 1) \\ &= \Pr(2^{\sum_{i \text{ even } \leq m} R_i 2^{-\sum_{j \text{ odd } < i} R_j}} \geq 2^{a_0-1}) \\ &\leq \frac{\mathbb{E}[2^{\sum_{i \text{ even } \leq m} R_i 2^{-\sum_{j \text{ odd } < i} R_j}]}{2^{a_0-1}}, \end{aligned} \quad (145)$$

where the last step follows from the Markov inequality. The idea is now to provide an upper bound on the quantity $\mathbb{E}[2^{\sum_{i \text{ even } \leq m} R_i 2^{-\sum_{j \text{ odd } < i} R_j}]$. Let $X = \sum_{i \text{ even } \leq m} R_i 2^{-\sum_{j \text{ odd } < i} R_j}$. We have

$$\begin{aligned} & \mathbb{E}[2^X] \\ &= \sum_{l=1}^{\infty} \Pr(R_2 = l) \mathbb{E}[2^X | R_2 = l] \\ &\stackrel{(a)}{=} \sum_{l=1}^{\infty} \frac{1}{2^l} \mathbb{E}[2^X | R_2 = l] \\ &= \sum_{l=1}^{\infty} \frac{1}{2^l} \mathbb{E}[2^{\frac{R_1}{2^l}}] \mathbb{E}[2^{\frac{X}{2^l}}] \\ &= \sum_{l=1}^{\infty} \frac{1}{2^l (2^{1-\frac{1}{2^l}})} \mathbb{E}[2^{\frac{X}{2^l}}] \\ &\stackrel{(b)}{\leq} \sum_{l=1}^{\infty} \frac{1}{2^l (2^{1-\frac{1}{2^l}})} (\mathbb{E}[2^X])^{\frac{1}{2^l}}, \end{aligned}$$

where (a) follows from the fact that R_i s are iid and X is self-similar and (b) follows from Jensen inequality. As a result, an upper bound on the quantity $\mathbb{E}[2^X]$ can be derived as follows. We have

$$\begin{aligned} \mathbb{E}[2^X] &\leq \frac{1}{2(2^{\frac{1}{2}}-1)} (\mathbb{E}[2^X])^{\frac{1}{2}} + \frac{1}{4(2^{\frac{3}{4}}-1)} (\mathbb{E}[2^X])^{\frac{1}{4}} \\ &\quad + \frac{1}{4(2^{\frac{7}{8}}-1)} (\mathbb{E}[2^X])^{\frac{1}{8}}. \end{aligned}$$

The equation $y = \frac{1}{2(2^{\frac{1}{2}}-1)} y^{\frac{1}{2}} + \frac{1}{4(2^{\frac{3}{4}}-1)} y^{\frac{1}{4}} + \frac{1}{4(2^{\frac{7}{8}}-1)} y^{\frac{1}{8}}$ has only one real valued solution y^* , and $y^* \leq 3$ (more precisely, $y^* \approx 2.87$). As a result, we have $\mathbb{E}[2^X] \leq y^* \leq 3$. Thus by (145) we obtain

$$\Pr(a_0 - \sum_{i \text{ even } \leq m} R_i 2^{-\sum_{j \text{ odd } < i} R_j} \leq 1) \leq \frac{3}{2^{a_0-1}}$$

Thus, given that $B_1 = 1$, we have:

$$\Pr(A_n \geq 2^{\sum_{i=1}^n B_i}) \geq 1 - \frac{3}{2^{a_0-1}}.$$

Or more precisely we have

$$\Pr(A_n \geq 2^{\sum_{i=1}^n B_i} | B_1 = 1) \geq 1 - \frac{3}{2^{a_0-1}}.$$

Now consider the case $B_1 = 0$. We show that a similar bound applies for A_n . Firstly, note that by fixing the value of n the distribution of R_1 is as follows: $\Pr(R_i) = \frac{1}{2^i}$ for $1 \leq i \leq n-1$ and $\Pr(R_1 = n) = \frac{1}{2^{n-1}}$. We have

$$\Pr(A_n \geq 2^{\sum_{i=1}^n B_i} | B_1 = 0)$$

$$\begin{aligned} &= \sum_{i=1}^n \Pr(A_n \geq 2^{\sum_{i=1}^n B_i} | R_1 = i, B_1 = 0) \Pr(R_1 = i | B_1 = 0) \\ &= \sum_{i \leq a_0-1, i \leq n} \Pr(A_n \geq 2^{\sum_{i=1}^n B_i} | R_1 = i, B_1 = 0) \Pr(R_1 = i | B_1 = 0) \\ &\quad + \sum_{i > a_0-1, i \leq n} \Pr(R_1 = i | B_1 = 0) \\ &\leq \sum_{i \leq a_0-1, i \leq n} \frac{1}{2^i} \frac{3}{2^{a_0-1-i}} + \frac{2}{2^{a_0-1}} \\ &\leq \frac{3a_0}{2^{a_0-1}}. \end{aligned}$$

Hence, considering the two cases together, we have:

$$\Pr(A_n \geq 2^{\sum_{i=1}^n B_i}) \geq 1 - \frac{3(1+a_0)}{2^{a_0}}.$$

Hence, the proof follows with $c_2 = 3$. \blacksquare

Lemma 21: Let $\alpha < 1$ be a constant. We have for $x \in (0, \frac{3}{4}]$

$$x \log\left(\frac{1}{x}\right) \leq c_3 (x(1-x))^\alpha, \quad (146)$$

where

$$c_3 = \frac{2}{(1-\alpha) \ln 2}. \quad (147)$$

Proof: By applying the function $\log(\cdot)$ to both sides of (146) and some further simplifications, the inequality (146) is equivalent to the following: For $x \in (0, \frac{3}{4}]$

$$\log\left(\log\left(\frac{1}{x}\right)\right) \leq \log c_3 + (1-\alpha) \log \frac{1}{x} + \alpha \log(1-x).$$

As $x \leq \frac{3}{4}$, we have $\alpha \log(1-x) \geq -\log 4$. Hence, in order for the above inequality to hold it is sufficient that for $x \in (0, \frac{3}{4}]$

$$\log\left(\log\left(\frac{1}{x}\right)\right) \leq \log \frac{c_3}{4} + (1-\alpha) \log \frac{1}{x}.$$

Now, by letting $u = \log \frac{1}{x}$, the last inequality becomes

$$(1-\alpha)u - \log u + \log \frac{c_3}{4} \geq 0, \quad (148)$$

for $u \geq \log(\frac{4}{3})$. It is now easy to check that by the choice of c_3 as in (147), the minimum of the above expression over the range $u \geq \log(\frac{4}{3})$ is always non-negative and hence the proof follows. \blacksquare